TWO ESSAYS IN FINANCE

BY

EUNJI OH

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Finance
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Doctoral Committee:

Professor Timothy C. Johnson, Chair
Associate Professor Jaewon Choi
Associate Professor Dana Kiku
Professor Louis K. C. Chan

Abstract

The first essay examines whether systematic equity risk of firms reflects the risk of their R&D strategies at various angles. More novel R&D strategy is risky because it can be related to more extreme outcome. This risk could indirectly affect the firm's systematic risk. In the case of success of the strategy, the productivity of technologies developed by novel R&D strategy could be procyclical; thus public firms with more novel technologies could be more subject to the aggregate risk. To investigate this problem, I devise an ex-ante measure for the novelty of innovation, Tech Synthesis Level (below TSL), which quantifies the degree that new technology is drawn from prior technologies in the far different technological fields, using patent citations. I find that patents with high TSL are associated with more extreme technological outcomes both at patent-level and startup-level. At public firm analyses, I find that high TSL is associated with high abnormal returns by 2.532 percent (annualized) and high systematic volatility. These findings support the hypothesis that a failure probability of the R&D project increases systematic risk. I also find evidence that high TSL patents are technologically more productive when aggregate innovation is very active, so firms with high TSL patents are subject to high systematic equity risk.

The second essay studies the effect of intangible collateral, which has gradually increased since the '90s, by testing hypotheses inspired by [Ai et al., 2018]'s collateralizability premium. Firms with more collateralizable capital have lower stock returns due to the insurance effect of the capital during economic recessions when financial constraints get tighter. If intangible collateral also can relax financial constraint, firms with intangible collateral are expected to have lower stock returns than the other similar firms without collateralizable intangible capital. I add empirical evidence by using *Dealscan* data of US-originated secured long-term loans. I find that firms using intangibles as collateral in addition to traditional collateralizable assets have higher stock returns than the other firms pledging only tangible assets to secure corporate loans. Also, they could achieve the

similar or even slightly higher level of leverage, implying intangible collateral also can relax financial constraint. This is not assumed possible in many theoretical and empirical studies. Even with matching analysis I find that firms pledging intangible capital as collateral still have higher stock returns than the other similar firms without intangible collateral. The empirical evidence I find does not fully support the collateralizability premium hypothesis.

Acknowledgments

I would like to express my deepest gratitude to my advisor Prof. Timothy C. Johnson for the endless support of my study and research, for his patience, motivation, and immense knowledge. His guidance inspired and helped me in all the time of research and writing of this dissertation. Also, I would like to thank the rest of my thesis committee: Prof. Jaewon Choi, Prof. Dana Kiku, and Prof. Louis K. C. Chan, for their insightful comments and encouragement. This dissertation would have never been possible without their help and guidance,. I also thank Prof. Deepak Somaya for his comments which helped me broaden the perspective of my study. Also, I thank Prof. George Pennacchi and Prof. Heitor Almeida for their support as PhD program director. In addition, I thank Ms. Yolanda Jackson and Ms. Denise Madden for their administrative help. I thank my fellow PhD students and graduates for the productive discussions, and for all the fun we have had in the last years in Champaign-Urbana. Last but not the least, I would like to thank my family and friends for their love and spiritual support.

Table of Contents

List of Tables

List of Figures

Chapter 1

Creation by Connecting the Dots: The Risks and Rewards of Novel Innovation

1.1 Introduction

This paper examines whether the systematic equity risk of firms reflects the risk of their R&D strategies. Previous studies have found answers to this question. Some studies show that the quantitative aspects of innovation matter for stock returns by using proxy variables such as the amount of R&D expenses or the intensity of R&D activities. For example, [Hirshleifer et al., 2013] argues that innovation efficiency, which is measured by the standardized number of forward citations that a firm's patents receive, predicts the cross-section of stock returns. However, there are fewer studies in the finance literature that discuss how the qualitative aspects of firm-level innovation affect the firm and stock returns. When considering a firm as a collection of R&D projects, it is reasonable to question whether the qualitative dimension of R&D projects, in other words, the R&D strategy of a firm, affects firm-level outcomes and, ultimately, the equity risk.

One representative quality of innovation is novelty. In a Harvard Business Review article, [Hopp et al., 2018] summarizes 40 years of research in disruptive innovation. They define disruptive innovation as the creation of new knowledge and the commercialization of completely *novel* ideas or products while incremental innovation is defined as the improvement of previous technology. The authors state that novel innovation focuses on long-term impact and is disruptive, bringing change in the relationship between customers and suppliers and creating unprecedented product categories. The authors state that in most cases, firms depend on advancements in technologies to achieve novel innovation.

In addition, [Ahuja and Lampert, 2001] study how large established firms' performs breakthrough innovations. They define novel technologies as new and unfamiliar technologies to firms in which the firms do not have enough prior experiences or expertise. Therefore researching and developing novel technologies requires significant costs of learning new knowledge. They also find that the exploration of novel technologies for large firms is positively associated with breakthrough innovation. Thus, novel technologies are expensive to develop but bring technological breakthrough and consequent economic success. However, not all novel technologies are related to breakthrough innovation or success, it could be subject to higher variability in the outcome ([Fleming, 2001]). Also, even in the case of success, the novel innovation may take a longer time to be fully realized ([Verhoeven et al., 2016]).

Therefore, it is plausible to expect that novel innovation is associated with greater growth opportunity through higher technological and economic impacts conditional on its success. At the same time, novel innovation can have a greater likelihood of failure in R&D process because it is difficult to develop the disruptive technology. Also, it could be more exposed to risk of being obsolete when the market is not ready to adopt new innovation. Additionally, novel innovation may entail higher costs than non-novel innovation due to its technological difficulty as in [Ahuja and Lampert, 2001] and [Verhoeven et al., 2016]. In summary, novel innovation may be related to extreme outcomes, a greater likelihood of failure, and a more significant technological impact because the innovation is used by the associated field's technologies and the technologies of other fields conditional on its development completion(success).

Considering the innate riskiness of novel innovation, it is plausible to assume that there are some differences between firms pursuing novel innovation and firms pursuing incremental innovation while viewing a firm as a collection of R&D projects. In this paper, I empirically study the effect of the novelty of innovation on various patent outcomes, early-stage firms, and public firms. By putting together all the evidence from the analyses, I find channels where the extent of novelty of an R&D strategy affects systematic equity risk in the cross-section of stock returns.

How can novelty of innovation be defined and measured precisely with existing data? What types of innovation are novel? Steve Jobs, the former CEO of Apple, gave his opinion on creativity in an interview with *Wired* magazine([Wolf, 1996]) as in the following quote:

> "*Creativity is just connecting things. When you ask creative people how they did something, they feel a little guilty because they didn't really do it, they just saw something. It seemed obvious to them after a while.*"

Opinions similar to that of Jobs exist whereby true innovation is believed to stem from the unique recombination of existing things ([Dahlin and Behrens, 2005],[Kaplan and Vakili, 2015],[Verhoeven et al., 2016]). Borrowing this idea, I define novelty of innovation as the creation of new technology by connecting prior technologies in far different fields from the created technology. Therefore, the critical quality to quantify is the technological distance in the technology space. The difference between technologies can be measured by the distance between two technologies in the technology space. To compute this distance, I use patents' backward citations to compute technological distance empirically. Because all US patents must disclose information material to patentability that is relevant to prior patents, almost all patents have backward citations. Therefore, this information is an excellent source of knowledge for the computing of technological distance for almost all patents. I define the average technological distance *Tech Synthesis Level* (*TSL*).

For example, if a medical diagnostic technology is built upon civil engineering technologies, the medical technology is considered to be novel because medical technologies does not have much overlap with civil engineering technology. Also, one can think of an analogy with journal papers. If a finance paper draws new financial knowledge mostly from psychology papers and neuroscience papers rather than finance papers, it is considered very novel finance paper with the novelty defined above. In other words, the novelty of innovation defined above is also unique recombination of prior technologies that has not been tried by many people.

My main findings are as follows[1]: At patent-level, I find that higher *TSL* is associated with a greater likelihood of winning an innovative product award and also a greater likelihood of failing to maintain the patent(early expiration). These findings imply that *TSL* is an ex-ante proxy for the risk associated with outcomes. In other words, a patent that uses outside-the-box technologies is more likely to meet with either huge success or huge failure.

In addition, by using a novel dataset from Crunchbase, I find that startups pursuing more novel innovation represented by high *TSL* are more likely to go public by running multinomial logit regression with various exit types of startups on *TSL*. These startups require higher investments (funding amounts) for research and to develop the technologies. These findings support the hypothesis that high *TSL* technology could bring highly successful outcomes at the firm level but requires more costs

[1]For detailed description of empirical methodology, please refer to Section 1.3

and thus have high chance of failure. On the other hand, *TSL* has no significant relationship with other types of startup exits such as being acquired by larger private or public firms. If one views that being acquired implies a deterministic outcome to a venture while an IPO implies an uncertain (risky) future outcome, it is reasonable to conclude that pursuing a high *TSL* R&D strategy for a venture is related to a risky outcome than deterministic outcome.

Does this effect of *TSL* aggregate at public firm level? [Hirshleifer et al., 2018] shows that Innovative Originality is one technological characteristic that is aggregated at firm level and has some effects on stock returns, defining Innovative Originality as a measure for breadth of knowledge inputs. Similarly, I find that a technological distance represented by *TSL* is also aggregated at public firm level and has a positive association with stock returns and systematic risk. The technological breadth and distance are two major aspects of the technological complexion of both an invention's roots and its impact([Jaffe and de Rassenfosse,]). The breadth cares about the variety of the knowledge inputs for an invention, while the distance cares about the closeness between the invention and its knowledge inputs. For example [Trajtenberg et al., 1997] studies how the technological breadth(originality) and distance are related to technological impact.

For the firm-level analyses, I find that higher *TSL* is associated with (1) higher firm-level returns using Fama–MacBeth regression and a high *TSL* portfolio has higher abnormal returns of 2.532 percent (annualized) compared to a low *TSL* portfolio. (2) An increase in future return volatility (total and systematic) and ROA volatility but no increase in idiosyncratic volatility with panel regression. The findings imply a possibility that *TSL* is priced in the cross-section of stock returns subject to some type of systematic risk source. Alternatively, the findings support a meaningful association between *TSL* and systematic equity risk.

I suggest a possible channel that explain the positive association between *TSL* and systematic equity risk. First, firms with high *TSL* can have high co-movement with aggregate innovation productivity because novel innovation is related to higher technological impact; novel technology is used by more post-technologies in various fields conditional on its successful development. If this is the case, and innovation productivity risk carries the positive price of risk, firms with high *TSL* have a higher risk premium. I find indirect evidence that high *TSL* patents are more productive when the aggregate innovation level is high. I also support the channel by showing that *TSL* factor series is positively correlated with aggregate knowledge capital productivity shock, which has positive correlation around 0.3 with aggregate consumption shock. In addition, knowledge capital productivity risk has a positive slope coefficient in the cross-sectional regression with various test asset portfolios. Finally, when firm-level beta to intangible capital productivity shock is included in the Fama–MacBeth regression, *TSL*'s explanatory power becomes statistically insignificant, which implies that the effect of *TSL* is subsumed by knowledge capital productivity risk. This implies that TSL captures the systematic risk associated with knowledge capital productivity.

Of course alternative explanations for why *TSL* should be positively related to firm risk and stock returns are also possible. For example, *TSL* could be one of the idiosyncratic risks but it can affect systematic risk indirectly suggested in [Berk et al., 2004]. The failure risk associated with novel R&D strategy can indirectly change the risk premium. [Berk et al., 2004] show that the failure risk of an R&D project can indirectly increase systematic risk in their theoretical model. They view a firm as a multi-stage R&D project. The firm has option to invest in the next stage of R&D project at the end of each R&D stage, and realizes future cash flows upon its completion of the multi-stage R&D project. Thus the firm is considered as a compound option on future cash flows. The failure risk of each R&D stage increases the threshold of an option to invest in the next-stage R&D project. If the investment is not made, a firm as a compound option on future cash flows has a higher risk premium due to unresolved uncertainty. Thus, failure risk can indirectly increase systematic risk. The finding in this paper that high *TSL* is associated with a higher likelihood of patent failure and high stock returns supports the model in [Berk et al., 2004], because a patent's failure which is an early expiration is considered as happening in a new product/process development stage.

This paper contributes to the literature on innovation and the cross-section of stock returns. To my knowledge, that this is the first paper that reveals a relationship between technological distance and various innovation outcomes including firm dynamics and the cross-section of stock returns. Additionally, this paper provides evidence that the association stems from systematic risk.

This paper also contributes to the innovation literature by suggesting a new novelty measure for innovation, which shows the extent to which a new patented technology is built upon extremely different prior technologies and its association with more extreme outcomes. The novelty measure, *TSL*, could be useful in evaluating technological impact or riskiness of R&D strategy ex-ante to not only investors but also corporate management.

The paper is organized as follows. Section 1.2 presents a literature review. Section 1.3 explains the empirical methodology used in the analyses in detail. Section 1.4 describes the data used in the paper. Section 1.5 presents the findings with a discussion, and Section 1.7 concludes.

1.2 Literature Review

There has been long-standing debate as to the definition of novel innovation and how it is related to technological and economic outcomes. As stated in Section 1.1, one opinion is that novel innovation is a result of unexpected re-combination. Others argue that novel innovation is creation from almost nothing, occurs typically near the beginning of the technological trajectory, and is highly original([Trajtenberg et al., 1997]).

[Dahlin and Behrens, 2005] states that novel invention should satisfy two conditions. First, a novel invention recombines prior arts in a new and different way. Second, many patents issued after a novel invention should imitate the novel invention's way of recombining prior technologies. The authors find this hypothesis to be true at least within the domain of tennis racket technology.

[Verhoeven et al., 2016] define novel technology in a similar way. The authors argue that a technology is novel in two dimensions, (1) novelty in recombination and (2) novelty in knowledge origins. The authors find that technological novelty is positively related to the variance of the technological impact.

This discussion exists in other research domains other than innovation. In 2013, [Uzzi et al., 2013] published an interesting article about atypical combinations of journal disciplines and their scientific impact in *Science*. The authors argue that balancing atypical knowledge with conventional knowledge may be the most important factor that links a study's novelty and its scholarly impact in the future.

In this paper, I focus on recombination-type innovation because it is relatively easily measured with empirical data. There are far fewer patents that do not have any backward references. The majority of these are probably a result of missing information, and not a lack of reference to any prior art, since US patent law clearly states the obligation to disclose material references. Although there is no clear definition in the innovation literature for a truly novel innovation, whether it be recombination or creation from a zero-base, I add evidence that recombination-type innovation could be novel and technologically valuable.

In terms of measuring *TSL*, I use technological distance between technological classes. A well-known study by [Jaffe, 1989] is the first paper to attempt to measure technological closeness between two firms' patent portfolios by using cosine similarity. If two firms have a similar distribution of technological classes in their portfolios, they are deemed to be close technologically. Inspired by this paper, I measure technological class-level closeness using cosine similarity and use it to calculate patent-level tech synthesis degree.

There are various papers that discuss optimal measures for technological distance. For example, [Younge and Kuhn, 2016] devise a text-based vector space model to measure patent-to-patent similarity. This captures previously unseen similarity between patents under a discrete technological classification system and requires a machine-learning technique.

[Yan and Luo, 2017] compares 12 technological distance measures to identify one that represents technological network consistently over time and has high correlation with other measures. The authors conclude that a normalized co-reference-based measure has high correlation with other patent network maps drawn from different distance measures and is also highly consistent over time. The authors find that class-to-class cosine similarity performs worse than co-reference-based distance. For the purposes of technological network mapping and analysis, it may be necessary to follow [Yan and Luo, 2017].

However, based on my analysis, while not reported, co-reference-based distance is extremely skewed and, thus, it is difficult to show variability across pairs of technological classes. Cosine similarity-based distance is still highly skewed but less so than co-reference-based distance. This fact is also reported in [Yan and Luo, 2017]. Thus, I choose conventional cosine similarity to measure technological distance between two patent classes.

This paper contributes to the literature on the relationship between innovation and the cross-section of stock returns. There are several papers that find that R&D intensity and efficiency matter for stock returns([Hirshleifer et al., 2013],[Chan et al., 2002]), but they focus on quantitative aspects of innovation not qualitative aspects.

A recent paper by [Hirshleifer et al., 2018] focuses on the qualitative aspects of innovation, which they call Innovative Originality. The authors study the relationship between Innovative Originality and the cross-section of stock returns in a behavioral economics sense. Firms with high Innovative Originality are difficult to value and are subject to high valuation uncertainty. This can cause firms to be undervalued and, consequently, to have higher stock returns. The authors also found that Innovative Originality predicts higher, more persistent, and less volatile profitability. Innovative Originality is defined as the total number of unique technological classes of patents to which a focal patent refers. Thus, Innovative Originality measures the breadth of the knowledge used in a patent while *TSL* focuses on the degree of difference between the originating patent and its backward references. [Hirshleifer et al., 2018] does not use technological distance information between the originating patent and its backward citations. Therefore, it is clear that what *TSL* measures is different from Innovative Originality. The difference between the technological distance and the originality is also documented in [Trajtenberg et al., 1997]. This current paper is also distinguished from [Hirshleifer et al., 2018] because I explain the cross-sectional difference using a systematic risk story.

There are a strand of papers that innovation is systematically risky. For example, [Hsu, 2009] empirically shows that technological innovation increase expected stock returns and premiums at the aggregate level by showing that shocks to innovation measured by patent flow and R&D flow are positively related to aggregate expected stock returns. The pattern is demonstrated internationally. In this paper, I find that novel technological innovation productivity is pro-cyclical and systematically risky, consistent with the findings in [Hsu, 2009]. I use forward citations as a measure of innovation productivity while [Hsu, 2009] uses the number of patents and R&D stock as a measure of innovation productivity.

There is another recent study on the technological closeness of two firms and the cross-predictability of stock returns by [Lee et al., 2018]. This study uses classic measures of technological closeness between two firms' patent portfolios to determine technologically close firms. The authors' channel leans more toward behavioral explanations, namely, that return predictability stems from slow price adjustment to more nuanced technological news transmitted via the technological link. The technological peer firm's stock return predicts the originating firm's stock return after controlling for industry fixed effect. This effect is stronger for firms with intense and specific technology focus with limited investor attention measured by analyst coverage. This paper studies the effect of information transfer through technological links but not the effect of technological characteristics on a firm's stock returns.

This paper also contributes to the literature concerning how to measure the value of innovation. For example, [Kogan et al., 2017] suggests an event study as a way to measure the private value of individual patents. *TSL* could be an ex-ante measure of technological value and impact, and it can be used together with the private value measure in [Kogan et al., 2017] to assess the value of patents from various perspectives.

1.3 Empirical Strategy

To find evidence supporting the hypothesis that the characteristics of an R&D strategy affect systematic risk and stock returns (see section 1.1), I exhaustively investigate how *TSL* is related to outcomes at three different levels—patents, early-stage ventures, and public firms. First, I explain how to compute the novelty measure, *TSL* , in detail. Next, I explain how I determine whether the *TSL* is an appropriate measure for the riskiness of novel innovation with patent-level analysis. Then, I identify the association between a firm's R&D strategy (firm-level TSL) and firm outcomes and stock returns. Finally, I explain an empirical strategy that determines systematic risk channels to explain the empirical findings above.

1.3.1 The Novelty Measure: *TSL*

To measure the novelty of innovation in terms of technological distance, I first measure how far a pair of technological classes are from each other using backward citations of US patents. Backward citations are the prior arts that are cited by a patent. I use *CosineSimilarity* to measure the class level closeness and then subtract it from 1 to make it a proper distance measure. The cosine similarity measures how similar the backward citations distributions over all technological classes are between two technological classes. Simply, *CosineSimilarity* is a correlation coefficient between two distributions of backward citations. Thus, *CosineSimilarity* lies between 0 and 1. If a pair of technological classes have exactly the same backward citation distribution, the cosine similarity between the two classes is 1. If, by contrast, the backward citations profiles do not have any technological classes in common, the cosine similarity between the two classes should have a value of 0. Using equation (1.1) and equation (1.2), I measure technological distances between a patent in class i and its backward-referenced prior arts in class j for each year in the sample. Since the number of backward citations from class i to class j varies over time, technological distance between the two classes also varies over time. I count backward citations from class i to class j from the beginning of the sample period (year 1976) to year t and take it as $C_{ij,t}$ in equation(1.1)[2]. Consequently, the technological distance measure varies over time.

$$CosineSimilarity_{ij,t} = \frac{\sum_{k=1}^{n} C_{ik,t} * C_{jk,t}}{\sqrt{\sum_{k=1}^{n} C_{ik,t}^2 * \sum_{k=1}^{n} C_{jk,t}^2}} \tag{1.1}$$

$$TechnologicalDistacne_{ij,t} = 1 - CosineSimilarity_{ij,t} \tag{1.2}$$

where $C_{ik,t}$: the number of backward citations of class i to class k from 1976 to year t

$C_{jk,t}$: the number of backward citations of class j to class k from 1976 to year t

Next, I take the mean of $TechnologicalDistance_{ij,t-1}$ over all backward citations of an originating patent p in class i to be the technological distance (*TSL*) of the patent p issued in year t (equation(1.3)). To remove technological class-fixed effect, I subtract the average $TechnologicalDistance$ of class i. This measure is similar to a weighted average of the backward citations. A high number implies that the originating patent is based upon very different prior patents. After removing the class-level mean effect, *TSL* can have a value between -1 to 1.

$$TSL_{p,i} = \frac{1}{J}\sum_{j=1}^{J} TechnologicalDistance_{ij} - Avg.TSL_i \quad \forall j = 1, 2, \ldots, J \tag{1.3}$$

Figure 1.1 shows the distribution of TSL_p. It is centered around zero, which implies that the majority of the patents have *TSL* of its class mean. If a patent lies on the right tail, the patent refers to starkly different prior patents on average. For an analogy, patents lying on the right tail are similar to those in finance papers that refer mostly to biology and psychology journals rather

[2] The reason I only count the number of citations from the year 1976 is limited data. The US Patent and Trademark Office (USPTO) only provides complete citations data for patents issued from the year 1976.

than finance or economics journals.

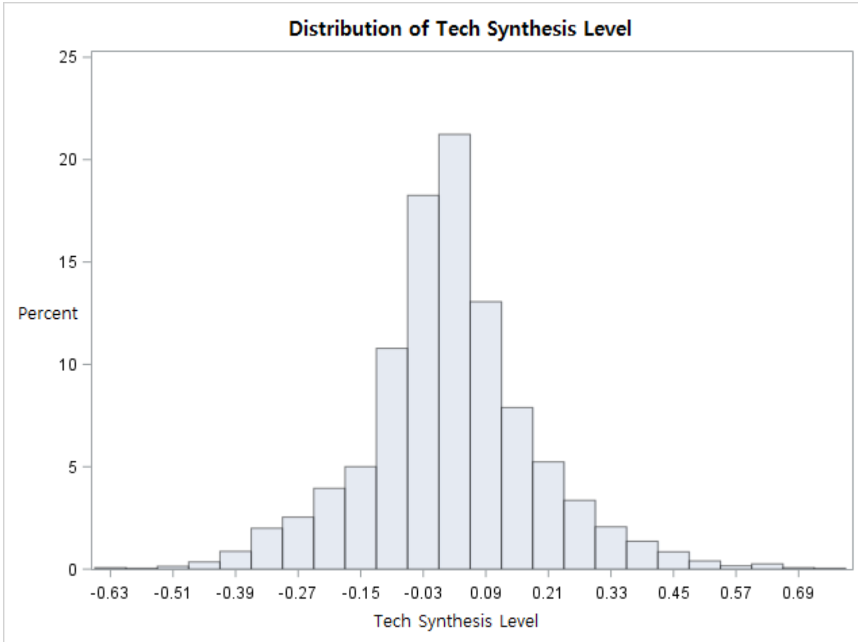

Figure 1.1: ***TSL* Histogram (Patents issued from 1976 to 2010)**

1.3.2 Valildating *TSL* as a technological risk measure

I first show that TSL predicts extreme technological outcomes at the patent level. In this way, I can plausibly say that TSL captures the riskiness of novel R&D strategy and whether it brings success or failure. I use winning an R&D award as a proxy for a highly successful technological outcome. I obtain a list of products that received an R&D 100 award and match it with relevant patents used in the product development. As I discuss in section 1.4, R&D100 awards are given to the top 100 revolutionary technologies of the past year and are often referred to as the *Oscars of invention.* I first identify patents related to awarded products and assign a value of 1 to an indicator variable representing the receiving of an award. I run a logit regression of the indicator variable on *TSL* and other control variables. This shows if TSL positively predicts the likelihood of receiving an award, which represents a successful technological outcome.

$$Pr(I_{award,p,i,t} = 1) = \frac{1}{1 + \exp(\alpha + \beta_1 TSL_p + \Sigma\beta_x Controls_p + YearFE)} \tag{1.4}$$

Next, I choose an early expiration of a patent due to failure to pay a maintenance fee as a proxy for an extremely negative outcome. The US patent system requires mandatory renewal fees for patents at the end of the fourth, eighth, and 12th year. If these fees are not paid, the patent expires right away. These fees for a patent are in the range of several hundred dollars for individuals to around eight thousand dollars for firms[3]. From the perspective of an entire R&D project, issued patents are still be used in development of new products or processes. Therefore, early expiration of a patent can be considered as a failure in development. To determine if high TSL patents are more likely to fail, I run a logit regression of an indicator variable for patent expiration on TSL and other control variables. The dependent variable has value of 1 if a patent fails to pay a maintenance fee and the patent is expired(equation(1.5)). I include a patent's economic value as a control variable to see if the reason behind early patent expiration is its economic value rather than technological novelty.

$$Pr(I_{expire,p,i,t} = 1) = \frac{1}{1 + \exp(\alpha + \beta_1 TSL_p + \Sigma\beta_x Controls_p + YearFE)} \tag{1.5}$$

I also investigate if TSL is related to more extreme outcomes at the firm level, particularly with a startup sample. Although the ultimate goal of this paper is to provide empirical evidence of the positive relationship between TSL and stock returns with a public firm sample, the marginal impact of a patent is not economically significant to many public firms that have huge patent portfolios. The relationships with patent-level *TSL* would be more noticeable with a sample that is composed of firms that are more affected by one more patent and its qualitative characteristic. Startups fit well with this condition because they typically have no patent or a small number of patents. According to [Dalle et al., 2017], only 50,000 firms among 447,000 listed in *CrunchBase*[4] own one or more patents. This is slightly more than 10 percent. In other words, around 90 percent of startups worldwide do not have any patent at all. Thus, adding one issued patent to an empty portfolio should have a greater significant impact for startups compared to publicly listed firms that add one more patent to their huge patent portfolio. I admit that startups and public firms do not share many common characteristics, but I believe investigating the startups could help to explain how novel innovation is related to firm risk and the extent of the effect.

The biggest limitation problem is that startups do not have any stock return data. Instead, I

[3]The numbers are based on a 2014 USPTO maintenance fee system

[4]a commercial database providing startups' basic and funding information

consider the likelihood that going public changes as *TSL* increases. For startups, an IPO exit is the greatest success. Therefore, If high *TSL* is truly associated with an extremely successful outcome, it should increase the likelihood of going public for startups. To check the relationship between *TSL* and the downside potential, I use unsuccessful exits("Closed") of startups as a proxy for failure. I run multi-logit regression with various types of startup exits as a dependent variable(equation(1.6)). I set the default case as "Not Having Exited the Startup Market Yet." The multi-logit regression shows how much more likely it is that the firm will go public than the default case scenario. Additionally, I include other types of startup-exit such as "Being Acquired by Other Firms" or "Expanding by Making Acquisitions." All the coefficients show the relative propensity of the alternative scenario to the default scenario.

$$Pr(Y_{exit,i} = n) = \frac{\exp(\mathbf{X_i}\beta^{(n)})}{\Sigma_{k=0}^{4}\exp(\mathbf{X_i}\beta^{(k)})} \qquad \forall n = 0, 1, .., 4 \tag{1.6}$$

where n=0: Not having exited (default scenario), $\beta^{(0)} = \mathbf{0}$

n=1: Closed (unsuccessful exit)

n=2: IPO (most successful exit)

n=3: Was acquired by other firms (less successful exit)

n=4: Made acquisitions (success without exit)

I also test if high *TSL* entails more costs. High research and development costs definitely change the NPV of projects and increase the likelihood of failure. Consequently, high costs raise the threshold for exercising the option to invest in next-stage R&D in the [Berk et al., 2004] model. To my knowledge, there is no way to obtain financial statement information on startups unless they go public. Therefore, I use the total funding amount as the total costs of investment assuming startups raise funding to expense all the money in new investments, particularly for R&D. Then, I run a regression on the total funding amount for *TSL* with other control variables. To be consistent with the hypothesis, high TSL should be associated with a high funding amount.

$$Y_i = \beta_0 + \beta_1 \frac{1}{T} \Sigma_{t=0}^{T} TSL_{i,t} + Controls + \epsilon_i \tag{1.7}$$

$$Y_i = \Sigma_{t=0}^{T} FundingAmount_{i,t} \tag{1.8}$$

T = The year of exit, if a startup has not exited then T is set to 7

1.3.3 TSL and firm risks

To find a link between TSL and firm risk, I run panel regressions for 60-month future volatility of excess returns on TSL and other firm-level controls. In practice, the volatility of excess return is used as a simple risk measure for a stock. I choose 60-month monthly return volatility because innovations typically take time to be in effect. Although [Turan G. Bali, 2016] shows that the level and distribution of excess return volatilities are similar regardless of the window used, volatility based on daily returns within one month or three months would not reflect information on technological innovation compared to five-year monthly volatility. Specifically, if TSL is associated with any type of systematic risk, it should have positive relation with systematic volatility but not with idiosyncratic volatility. To address the concern for within-firm correlation between errors, I cluster standard errors by firm.

$$Vol_{i,t,t+5} = \beta_0 + \beta_1 TSL_{i,t} + Controls + YearFE + FirmFE + \epsilon_{i,t} \tag{1.9}$$

Additionally, it is plausible to investigate if profitability becomes more volatile as TSL increases. I use ROA as a profitability measure and run a panel regression for five-year ROA volatility on TSL and other control variables. The result of this regression reveals if TSL is positively associated with future ROA volatility.

Next, I run Fama–MacBeth regression to see if TSL predicts stock returns in the cross-section with various control variables. I form portfolios of stocks based on TSL and perform sorting analysis. If the HighTSL-minus-LowTSL portfolio has a positive and significant alpha controlling for conventional factors, the risks related to TSL should be different from the conventional factors. I explain empirical strategies to find out the new risk source in the next section 1.3.4.

1.3.4 The systematic risk channel

(1) The direct channel through knowledge capital productivity

I suggest that a systematic risk source related to novelty R&D strategy is one associated with knowledge capital productivity. Firms with high *TSL* can have high co-movement with aggregate innovation productivity because novel innovation (high TSL type) is related to higher technological applicability. This implies that novel innovation is used by more post-technologies in various fields conditional on its successful development. Thus it is likely that there are more number of firms or post technologies which requires the previous novel technology when overall economy is good and innovation is more active, on the other hand, there would be less number of firms or post technologies which utilize the previous novel technology in bad times when overall innovation is less active. Therefore, due to its positive association with technological applicability, the high-TSL type novel innovation can make a firm's innovation productivity fluctuate more with the aggregate economy-wide productivity. If this is the case, and innovation productivity risk carries the positive price of risk, firms with high *TSL* have a higher risk premium.

I first find evidence directly at the patent level by showing whether patent-level *TSL* is more productive when the overall economy-wide innovation is more active. A patent's technological productivity can be measured by the amount of forward citations it receives. If a patent is more cited, it is considered to have a bigger technological impact, and thus more productive scientifically and technologically. Also [Trajtenberg et al., 1997] show that technological distance of a patent is positively associated with greater applicability in terms of that the patent is cited by other patents in various fields. So I investigate if high *TSL* patents receive more citations when economy-wide innovation is very active so that more number of forward citations for all patents are made.

In the first step, I estimate the co-movement between an individual patent's productivity and aggregate innovation productivity, which is β_{fcites}. The high β_{fcites} of a patent shows that it receives more citations when overall economy-wide innovation is very active and thus more number of citations are made. The $\beta_{fcites}$is obtained from time series regression of an individual patent's forward citations on aggregate forward citations(equation(1.10)). I replace missing forward citations with zero. In this regression, I use relative forward citations that are adjusted for the technological class-level fixed effect. Then, I run cross-sectional regression of β_{fcites} on *TSL* to investigate if

TSL has positive predictability on β_{fcites}. If TSL turns out to be positively associated with β_{fcites}, it can be evidence for the hypothesis that TSL is positively related to *pro-cyclical* technological productivity.

$$Fcites_{p,t} = \beta_{0,p} + \beta_{fcites,p} AggregateFcites_t + \epsilon_{p,t} \tag{1.10}$$

$$\beta_{fcites,p} = \gamma_0 + \gamma_1 TSL_p + \gamma_2 \xi_p + Controls + \epsilon_p \tag{1.11}$$

Next, I test whether innovation productivity risk is one of the actual source of risk that is captured by TSL. If not only the patent-level TSL but the firm-level TSL is also positively associated with the aggregate innovation productivity, then a variable that captures a firm's exposure to the aggregate innovation productivity should subsume the explanatory power of TSL. Slightly different from the patent-level analysis above (equation(1.10) and (1.11)), I measure the aggregate innovation productivity by knowledge capital productivity of public firms. I test if the explanatory power of TSL disappears when including β for aggregate knowledge capital productivity in the cross-sectional regression of stock returns.

To complete this task, I first estimate aggregate knowledge capital productivity. I show that it has a positive price of risk so that high TSL firms have high stock returns due to greater exposure to knowledge capital productivity risk. Then, I estimate firm-level betas for the knowledge capital productivity shock. Finally, I run a Fama–MacBeth regression with the betas and check if TSL loses its statistical significance.

To estimate aggregate knowledge capital productivity, I borrow public firms' knowledge capital data from [Peters and Taylor, 2016]. Then, I run yearly cross-sectional regressions of firm-level revenue(SALE) on tangible capital(PPENT), knowledge capital, other intangible capital year by year to obtain a time series of knowledge capital productivity. All output and capital variables are divided by number of employees(EMP), so the productivity is per-capita value as in equation 1.12.

All capital variables are adjusted by relevant deflators at 2009 level.[5]

$$\frac{y_{it}}{L_{it}} = \alpha + \beta_{1,t}\frac{K_{tan,it}}{L_{it}} + \beta_{2,t}\frac{K_{know,it}}{L_{it}} + \beta_{3,t}\frac{K_{int,it}}{L_{it}} + \epsilon_{it} \tag{1.12}$$

$$\begin{aligned} \text{where } K_{tan,it} &= \text{Tangible capital of firm i in year t} \\ K_{know,it} &= \text{Knowledge capital of firm i in year t} \\ K_{int,it} &= \text{Other intangible capital of firm i in year t} \\ L_{it} &= \text{The number of employees of firm i in year t} \end{aligned}$$

Each beta in equation(1.12) represents capital-specific productivity. For example, $\beta_{2,t}$ is knowledge capital productivity in year t. Once I obtain the time series of capital-specific productivities, I estimate exogenous shock for each productivity with the VAR(1) model(equation(1.13)). I estimate consumption shock from BEA consumption series with the AR(1) model.

$$\begin{pmatrix} \beta_{1,t+1} \\ \beta_{2,t+1} \\ \beta_{3,t+1} \\ \beta_{4,t+1} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} + A_1 \begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \\ \beta_{4,t} \end{pmatrix} + \begin{pmatrix} \sigma_{1,t+1} \\ \sigma_{2,t+1} \\ \sigma_{3,t+1} \\ \sigma_{4,t+1} \end{pmatrix} \tag{1.13}$$

I use 130 test assets including 25 size-BM portfolios, 25 BM-investment portfolios, 25 size-investment portfolios, 25 size-profitability portfolios, 25 profitability-investment portfolios, and 5 TSL portfolios. I first estimate betas by running time-series regression and then running cross-sectional regression to determine if the slope coefficient of the knowledge capital productivity shock (λ_1 in equation(1.15)) is positive and significant. I run the time-series regression with entire series to estimate betas, and then run the cross-sectional regression of time-series average portfolio excess returns on estimated betas.

$$r^e_{p,t} = \alpha_p + \beta_{1,p}\sigma_{know,t} + \beta_{2,p}MKTRF_t + \beta_{3,p}HML_t + \beta_{4,p}SMB_t + \beta_{5,p}UMD + +\epsilon_{p,t} \tag{1.14}$$

$$r^e_{p,t} = \lambda_0 + \lambda_1\beta_{1,p} + \lambda_2\beta_{2,p} + \lambda_3\beta_{3,p} + \lambda_4\beta_{4,p} + \lambda_5\beta_{5,p} + \epsilon_{p,t} \tag{1.15}$$

[5] I use GDP deflator for SALE, intangible capital deflator for knowledge capital and other intangible capital, and physical capital deflator for tangible capital.

If the primary source of return difference between high TSL stocks and low TSL stocks is co-movement with aggregate knowledge capital productivity, adding β_{know} in the Fama–MacBeth regression should reduce the statistical power of TSL dramatically because β_{know} is a direct measure of co-movement between a stock return and aggregate knowledge capital productivity. I estimate β_{know} by running 60-month-rolling-window regression of a firm's stock returns on knowledge capital productivity shock from month t-60 to t and match the β_{know} with month t+1. Then I run regression in equation (1.16) to see if β_{know} subsumes TSL.

$$r^e_{i,t} = \gamma_0 + \gamma_1 TSL_i + \gamma_2 \beta_{know,i} + Controls_i + IndustryFE + \epsilon_{i,t} \tag{1.16}$$

(2) An alternative indirect channel

As in [Berk et al., 2004], various types of idiosyncratic risks associated with R&D could indirectly change a firm's systematic risk and risk premium. One of the risks is the failure risk associated with each R&D stage. Therefore, by providing empirical evidence that TSL is positively associated with the likelihood of failure and also with stock returns systematically, I can support the mechanism of the model in [Berk et al., 2004]. For the exercise validating TSL as the riskiness of the R&D project described in section 1.3.2, I satisfy the first condition to provide evidence of the association between TSL and the likelihood of failure. From the results of the analyses in section 1.3.3, I show whether TSL is associated with systematic risks and stock returns. Altogether, I provide empirical evidence supporting the hypothesis that the technological risks embedded in novel R&D projects can indirectly increase systematic risk and stock returns by hindering the resolution of uncertainty regarding whether to invest in the next stage of an R&D project.

1.4 Data

The sample period is from 1976 to 2010; thus, I can use public firm-patent maps provided by [Kogan et al., 2017]. Firm-years without any patents are excluded because it is not possible to compute TSL for these firm-years. For this reason, less than 1 percent of patents without any backward references[6] are excluded. Additionally, financial firms and utility firms are excluded based on SIC

[6] Backward references(citations) are the citations made by the originating patent to prior arts.

codes (4900-4999, 6000-6999). Small firms with market capitalization of less than 2 million dollars are excluded to reduce noise. Firms with negative book equity are excluded. To match financial data and patent data, I borrow the patent-PERMNO(CRSP) matching table from [Kogan et al., 2017], and COMPUSTAT book values are merged using the PERMNO-GVKEY link table. I perform extensive text-matching works to match novel datasets such as R&D 100 awards or Crunchbase with USPTO patents. I give detailed procedures in the relevant subsections below.

1.4.1 Financial Data

I use *CRSP* and *COMPUSTAT* to compute financial variables. CRSP data are from 1976 to 2010 while COMPUSTAT is from 1975 to 2009. All the book values are measured at the end of the fiscal year t and merged with year $t+1$ CRSP data. Market cap is measured by share price times the number of shares outstanding. The book value of equity is defined as shareholder equity (SEQ) plus deferred taxes($TXDITC$, if available), minus preferred stock($PSTK$). If SEQ is not available, I use common equity(CEQ) plus preferred equity($PSTK$) or assets(AT) minus liabilities(LT). If $PSTK$ is not available, I use redemption value($PSTKRV$) or liquidation value($PSTKL$). Market value of equity for the market-to-book ratio is measured at the end of the fiscal year. Intangible assets($INTAN$) are standardized by total assets to avoid bias from firm size. R&D expenses(XRD) are also standardized by total assets. Missing R&D expenses are zero only for firms that have ever had non-zero R&D expenses. Return on assets (ROA), a profitability measure, is measured by ordinary income before depreciation ($OIBDP$) divided by total assets(AT). Return on equity (ROE) has the same numerator as ROA but divided by fiscal year-end market capitalization. The definitions of gross profits and operating leverage follow [NovyMarx, 2013], [Novy-Marx, 2011] each. Stock betas are from WRDS beta suite, which is estimated using 60-month windows. I exclude firms if the number of observations used is less than 24 months. I obtained firm-level intangible capital data from [Peters and Taylor, 2016]. This data is available via WRDS. [Peters and Taylor, 2016] suggest a new way to measure intangible capital of public firms in the Compustat-CRSP universe. The authors decompose intangible capital into knowledge capital, organizational capital, and other intangible capital. They also provide total Q considering the effect of intangible capital. I also obtain consumption and GDP data from BEA.

1.4.2 Patent Data

USPTO provides complete information on US patents from the year 1976. Although firm-level data are available from the 1960s, because of the lack of patent information before 1976, the sample includes firm-years starting from 1976. I use CRSP PERMNO and USPTO patent number mapping data from [Kogan et al., 2017] for the sample period from 1976 to 2010. I also borrow patent-level economic value data from [Kogan et al., 2017], which is extracted from the three-day abnormal return around the patent issue date.

I use the USPC[7] technological class to compute class-level technological distance and the sample firms' *TSL*. I use all patents with at least one backward citation with complete class information to compute class-level technological distance. A patent's *TSL* measures average technological distance between an originating patent's class and the classes of cited patents. It is adjusted to be centered around the class's average distance. A firm's *TSL* is the average of the patent level measure across all issued patents in a given year. It is between -1 (within-class innovation) and 1 (outside-the-box innovation). Average forward citations measures how many times an originating patent receives citations from other patents compared to its technological class mean forward citations until the end of the sample period. If an originating patent has a value of 150%, this means that the patent receives 1.5 times more forward citations compared to its class-level mean forward citations. Thus, it is a little less affected by truncation bias because it is more like a cross-sectional measure after removing class-level mean effect.

All the bulk patent data are obtained from *Patentsview*[8].

The US patent system requires mandatory renewal fees for patents at the end of the fourth, eighth, and 12th year. If these fees are not paid, the patent expires right away. USPTO provides a dataset containing records of events related to maintenance fees. When an assignee fails to pay the fee, the event description variable is coded with "EXP." I compute year-lags between the patent issuance date and maintenance fee event date so that I can identify when a patent has expired.

[7]the US Patent Office's description states "The USPC is a system for organizing all US patent documents and many other technical documents into relatively small collections based on common subject matter." ([U.S. Patent Office, 2012])

[8]`http://www.patentsview.org/download/`

1.4.3 Breakthrough Innovation Data

I classify breakthrough invention by two different measures: the number of forward citations and the receipt of R&D awards. R&D100 awards are given to the top 100 revolutionary technologies of the past year, and they are often referred to as the *Oscars of invention*. To receive the award, a new technology or a product should be available for sale or licensing in the previous year. These data are already studied in the innovation literature([Fontana et al., 2013]), but there is still room for exploration. The history of R&D100 awards dates back to 1963, and the new significant products or processes of the past years are such as the digital wristwatch, antilock brakes, the automated teller machine, liquid crystal display, the halogen lamp, and the fax machine[9]. New inventions are judged by experienced panels composed of professional consultants, university faculty members, and industrial researchers[10]. The award winners include industry-leading corporations, public research organizations, and universities. Past winners include well-known corporations and universities, so the data is considered reliable as a reflection of truly novel inventions in the real world.

However, the award goes to products that could involve many different patents at the same time. To match R&D100-awarded products to relevant patents, I compare the titles of awarded products with those of patents, awarded organization names with patent assignee names, and the descriptions of awarded products with patent abstracts. For example, if a product receives an award in 2018, I match the awarded product with the awarded organization's patents issued in the past five years considering that R&D requires extensive time. For company name matching, I use a fuzzy string match algorithm. For abstract matching, I use a word-token based TF-IDF algorithm. The number of patents matched by awarded organization name and assignee name is 147,180. Among those patents, 5.9 percent are matched with awarded products by title, and 13.5 percent are matched with awarded products by either title or abstract.

1.4.4 Startup Data

I manually collected information on startups from *Crunchbase Pro*[11]. Crunchbase is a commercial database with information on innovative startups, and it was created in 2007. Basic information

[9] `https://hbswk.hbs.edu/archive/rd-magazine-online`

[10] `https://www.rd100conference.com/rd-100-award-judging-process/`

[11] For academic purposes, educational institutions can request free access to *Crunchbase* while individual users can subscribe to Pro for access to more detailed information

on firms includes organization, IPOs, M&As, and funding details. According to [Dalle et al., 2017], *Crunchbase* is updated daily, and users can add to and edit the database. As of September 2017, more than 3,000 global investment firms submit monthly portfolio updates to *Crunchbase* in exchange for free data access. Although the coverage of *Crunchbase* is not clearly defined, [Dalle et al., 2017]'s analysis shows that aggregate statistics on Venture Capital funding by country and year are similar to the numbers calculated from an alternative database, *OECD Entrepreneurship Financing Database*.

I only include startups founded in 2010 in the sample for several reasons. In 2010, the overall economy was recovering from the recent financial crisis; thus, startup foundations were possibly less affected by extreme economic conditions. Panel B of Figure 1 in [Dalle et al., 2017] shows the number of companies in the database by founding year, and more than 20,000 startups worldwide are reported in *Crunchbase* for 2010. Additionally, it has been over seven years since 2010, which is long enough to observe the successful or unsuccessful exits of these 20,000 startups. Therefore, I choose U.S. startups founded in 2010 for the analysis.

After collecting all relevant information on the startups founded in 2010, I match US patents with the startups by firm name and headquarter location. I first match them by names using a fuzzy string match algorithm and then compare country- and state-level locations to ensure proper matching.

There were 15,619 U.S. startups founded in 2010. Only 1,109 startups had at least one U.S. patent issued by May 2016. Among the matched startups, approximately 856 startups have information on funding amounts, exit status, and detailed investor information.

1.5 Findings and Discussion

1.5.1 Evidence from patents

Table 1.1 shows that high *TSL* is associated with a higher likelihood of receiving an R&D100 award(equation(1.4)). The columns (1), (2), and (3) use a matched sample on firm, title, and abstract. First-awarded firms are matched with assignee names in the patent database. Then, I narrow down the sample by matching awarded product titles with patent titles. I narrow down the sample further by matching awarded product descriptions with patent abstracts. This sample is from 1997 to 2010 due to the non-existence of product descriptions before the year 1997.

Columns (4), (5), and (6) are based on a sample that matches R&D100-awarded products and patents based on firm and title. Because the awarded product titles are provided from year 1963, I obtain the sample from 1976 to 2010.

Patents included in this analysis are only from the technological classes that include awarded patents. If there are no awarded patents in a technological class, the entire set of patents in the class is excluded from the analysis.

Based on columns (1), (2), and (3), a one-unit increase in *TSL* increases the likelihood of receiving an R&D100 award more than 50 percent relative to the likelihood of not receiving an award. In columns (4), (5), and (6), a one-unit increase in *TSL* increases the likelihood of receiving an R&D100 award by 18 percent compared to the probability of not receiving an award(equation(1.4)). I control for year fixed effect and cluster standard errors at the technological class level. Because the awards are given to the most innovative *"product"*, I believe this practice helps to control for unobservable variables regarding time and the possible correlation of errors within a technological class. In summary, these results imply that *TSL* proxies for technological novelty and value. Additionally, because R&D100 awards are given to commercialization-ready products, *TSL* could also capture the potential commercial value of a technology.

Table 1.2 contains the logit regression of patent expiration on *TSL* and other control variables. The dependent variable has a value of 1 if a patent has expired in the middle of its life. Patent expiration is a definite case of failure in the mid-stages of development which corresponds to the technological failure of an R&D stage in [Berk et al., 2004]. In columns (2) and (3) of Table 1.2, *TSL* positively predicts the likelihood of expiration. The likelihood of failing to pay a maintenance fee in the eighth year is approximately 1.1 times higher than the likelihood of keeping a patent alive. Even in the 12th year, the odd ratio is still slightly higher than 1, which implies that patents with high *TSL* are more likely to be abandoned before their natural termination.

However, *TSL* does not predict expiration in the fourth year as shown in column (1). This could be because firms are more likely to actively develop their services or products with patented technologies in the first three to four years. Thus, three years is not long enough for a firm to decide

Table 1.1: **Logit Regression: RND100 Awards and *TSL***

The sample period is from 1976 to 2010. A patent's *TSL* measures the average technological distance between an originating patent's class and the classes of cited patents. It is adjusted to be centered around the class's average distance. If a patent is matched with the RND100-awarded products, the RND100 indicator variable has a value of 1, otherwise 0. The number of backward citations is the total number of referenced patents that an originating patent has. The number of self-citations is the total number of backward citations made by the same assignee. The number of external citations is the total number of backward citations made by examiners. The top 5 percent of forward citations is an indicator variable if a patent has the top 5% forward citations relative to its technological class average. In columns (1), (2), and (3), awarded products are matched with patents by assignee name, product title, and abstract. Due to missing abstract information, the sample period for columns (1),(2), and (3) is from 1997 to 2010. In columns (4), (5), and (6), awarded products are matched with patents by assignee name and product title. Standard errors are clustered by the main technological class. The detailed empirical methodology is explained in Section 1.3.1.

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	RND100	RND100	RND100	RND100(T)	RND100(T)	RND100(T)
TSL	0.495**	0.458**	0.497**	0.179***	0.176***	0.176***
	(0.218)	(0.217)	(0.218)	(0.061)	(0.061)	(0.060)
Economic Value			8.47e-04			-0.0127*
			(7.59e-04)			(0.007)
No. Backward Citations	7.65e-04	1.33e-04	7.53e-04	7.82e-04	6.65e-04	8.61e-04
	(0.001)	(0.002)	(0.001)	(7.53e-04)	(7.65e-04)	(7.44e-04)
No. Self Citations	-0.166	-0.185	-0.161	0.212	0.211	0.193
	(0.281)	(0.279)	(0.280)	(0.165)	(0.165)	(0.160)
No. Examiner Citations	-0.801	-0.817	-0.769	0.177	0.175	-0.069
	(1.182)	(1.182)	(1.171)	(0.912)	(0.911)	(0.896)
Top 5% Fcites		0.745***			0.111**	
		(0.190)			(0.056)	
Constant	-2.762**	-2.778**	-2.794**	-2.660***	-2.666***	-2.604***
	(1.193)	(1.191)	(1.182)	(0.106)	(0.106)	(0.110)
Year FE	Yes	Yes	Yes	Yes	Yes	
Observations	64,171	64,171	64,171	135,582	135,582	135,582

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

whether to abandon patents.

With the results in Table 1.1, I empirically show that *TSL* is related to the riskiness of technological outcomes. The higher the *TSL*, the more extreme the outcome. A patented technology with fairly high *TSL* will either have a significant technological impact or be abandoned in the early stages of its life.

Another interesting finding shown in the table is that the economic value of a patent predicts the likelihood of expiration negatively. In other words, patents with positive stock market reactions are less likely to fail. This finding supports the economic value measure for a patent suggested by [Kogan et al., 2017] and also implicitly supports the view that stock market investors have the ability to assess fair values of new technologies.

1.5.2 Evidence from startups

I find firm-level evidence that high *TSL* is associated with pursuing uncertainty by examining the likelihood of an IPO from a startup. This also could provide evidence of a positive relation between *TSL* and extreme success, which is an IPO. Table 1.3 shows the multinomial logit regression results from the startup data. The dependent variable is a categorical variable. The default scenario is to operate a company as a startup in the market. *"IPO"* represents the exit of a startup by going public, *"Was Acquired"* represents the exit of a startup through an acquisition by another public or private firm, *"Made Acquisitions"* represents a startup that expands their business area by acquiring other startups or firms, and *"Closed"* represents the permanent exit of a startup from the market. This multinomial logit regression shows the relative likelihood of each scenario compared to the default case. The number of investors is the total number of individual and institutional investors in a startup. The total funding over the total equity funding ratio shows the extent to which a startup depends on equity financing. California is a dummy variable for startups headquartered in California. Startups are geographically clustered around California as are venture capitalists. Thus, being headquartered in California might affect the exit results. The ranking of total funding is based on the total funding amount up until the year 2016. This could also affect the exit results because some startups go public to access more funding in the public market rather than the private lending market. This is consistent with the hypothesis that startups are slightly more likely to go

Table 1.2: **Logit Regression: Patent Expirations and *TSL***

The sample period is from 1976 to 2010. A patent's *TSL* measures average technological distance between an originating patent's class and the classes of cited patents. It is adjusted to be centered around the class's average distance. The dependent variables are indicator variables, which have a value of 1 if a patent fails to pay the maintenance fee and is expired. Exp(4th) indicates whether a patent is expired in the fourth year after issuance. Exp(8th) indicates whether a patent is expired in the eighth year. Column(2) excludes patents expired before the eighth year. Exp(12th) indicates whether a patent is expired in the 12th year. Column(3) excludes patents expired before the 12th year. The dependent variable in column(4) has a value of 1 for patents expired in the fourth year or the eighth year. The dependent variable in column(5) has a value of 1 for patents expired in any year. The number of self-citations is the total number of backward citations made by the same assignee. The number of external citations is the total number of backward citations made by examiners. Standard errors are clustered by the main technological class. The detailed empirical methodology is explained in Section 1.3.1.

VARIABLES	(1) Exp(4th)	(2) Exp(8th)	(3) Exp(12th)	(4) Exp(4th or 8th)	(5) Exp(Any year)
TSL	0.058	0.098***	0.052***	0.088***	0.099***
	(0.035)	(0.029)	(0.020)	(0.030)	(0.030)
Economic Value	-0.001	-0.003*	-0.003***	-0.002	-0.003**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
No. Backward Citations	-0.005***	-0.003***	8.27e-05	-0.004***	-0.003***
	(0.001)	(0.001)	(3.20e-04)	(0.001)	(0.001)
No. Self-Citations	-0.171	0.125**	0.163***	0.0351	0.115*
	(0.122)	(0.053)	(0.032)	(0.073)	(0.061)
No. Examiner Citations	-0.223	-0.201*	0.144	-0.218*	-0.081
	(0.150)	(0.119)	(0.120)	(0.114)	(0.093)
Constant	-1.743***	-1.176***	-4.182***	-0.639***	-0.710***
	(0.151)	(0.121)	(0.129)	(0.111)	(0.0904)
Observations	1,110,203	1,013,296	988,079	1,110,203	1,110,203
Year FE	Yes	Yes	Yes	Yes	Yes

Robust standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

public when their *TSL* is higher compared to stay-as-usual ventures. Additionally, more investors decrease the probability of going public compared to the base case scenario. In addition, a higher funding amount (rank) increases the likelihood of a firm going public but to a lesser extent than *TSL*.

Being acquired by another firm could also proxy for the successful exit of a startup, but it is more noisy than an IPO exit because incumbent firms sometimes acquire startups to obtain ownership of their patents and technologies. In this case, an exit through acquisition is not as successful as an exit using an IPO. Therefore, it is not necessarily the case that a higher *TSL* predicts a greater likelihood of being acquired as *TSL* is related to either huge success or failure at the patent level. Additionally, [Wang, 2015] finds that large firms acquire startups from a similar technological field, and there are entrepreneurs who match their innovation to potential acquirers.

Column (3) in the table shows that only the equity funding ratio matters for exit through acquisition. *TSL* has no statistical power to predict the probability of being acquired. Startups with a lower proportion of equity funding are actually less likely to be acquired and more likely to stay as usual for a prolonged period. Column (4) shows that firms with many investors are more likely to acquire other startups than to stay as usual. *TSL* has no effect on the made-acquisition exit scenario. Column (1) shows no significant association between *TSL* and a higher propensity of unsuccessful exits. This could be because startups with more than one issued patent are already top-tier firms in the startup world. Therefore, it is less likely to see them unsuccessfully exit. They either go public or are acquired.

Table 1.4 shows how investment costs are related to *TSL*. The results in this table resemble contemporaneous correlation rather than a causal relationship. As mentioned in the previous sections, I assume the total funding amount as the total costs of researching and developing the technology. Consistent with the prediction, higher *TSL* is associated with a larger funding amount (costs). If *TSL* increases by 1 from -0.5 to 0.5, it increases total cost by $0.3 million. Although other variables, such as the number of investors or location, have far more significance statistically and economically than *TSL*, it is still meaningful to provide evidence that achieving high *TSL* is costly.

Together with Table 1.3, high *TSL* requires more investment but has a high chance of the firm successfully exiting via IPO. Therefore, achieving and maintaining a higher level of *TSL* can be

Table 1.3: **Multinomial Logit Regression: Startup Exits and *TSL***

The dependent variable is a categorical variable. The default scenario is to operate a company as usual as a startup in the market. *"IPO"* represents the exit of a startup by going public, *"Was Acquired"* represents the exist of a startup by being acquired by another public or private firms, *"Made Acquisitions"* represents a startup that expands its business area by acquiring other startups or firms, and *"Closed"* represents a startup that exits the market permanently. This multinomial logit regression shows the relative likelihood of each scenario compared to the base case. The number of investors is the total number of individual and institutional investors of a startup. The total funding over total equity funding ratio shows the extent to which a startup depends on equity financing. California is a dummy variable for startups headquartered in California. The total funding amount(Rank) is a ranking variable based on the total funding amount until 2016. The detailed empirical methodology is explained in Section 1.3.2.

VARIABLES	(1) Closed	(2) IPO	(3) Was Acquired	(4) Made Acquisitions
TSL	0.002	0.010***	0.002	0.004
	(0.007)	(0.004)	(0.003)	(0.003)
Number of Investors	-0.059	-0.073**	-0.003	0.084***
	(0.076)	(0.036)	(0.022)	(0.020)
Total Funding/Total Equity Funding	0.426	0.234	-2.183**	-0.246
	(0.400)	(0.517)	(1.045)	(0.796)
California	-0.131	-0.360	0.227	0.247
	(0.572)	(0.295)	(0.212)	(0.277)
Total Funding Amount (Rank)	1.40e-04	0.002***	7.96e-05	0.001***
	(2.60e-04)	(2.87e-04)	(9.84e-05)	(2.53e-04)
Constant	-4.469***	-9.125***	0.211	-7.918***
	(1.227)	(1.438)	(1.137)	(1.456)
Observations	856	856	856	856

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

either more profitable if everything goes well or can increase costs if operations do not proceed as expected compared to the case with a lower level of *TSL*. This result is consistent with patent-level evidence that *TSL* is associated with extreme outcomes, either huge success or huge failure.

Table 1.4: **OLS Regression: Startup Investment Costs and *TSL***

The dependent variable is the total funding amount raised between 2010 and 2016. The number of investors is the number of investors involved in funding. California is an indicator variable that shows if a startup is located in California. The detailed empirical methodology is explained in Section 1.3.2.

VARIABLES	(1) Total Funding Amount ($Mil)
TSL	0.310**
	(0.141)
Number of Investors	12.340***
	(0.922)
California	-43.750***
	(11.470)
Constant	13.610
	(9.250)
Observations	856
R-squared	0.172

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

1.5.3 Evidence from public firms

Table 1.5 shows the summary statistics of firms grouped by *TSL*. Firms are divided into three groups for *TSL* each year. The Low group has negative *TSL*, which means that they are engaged in more within-class innovations. On the other hand, The High group has positive *TSL*, which means that they are engaged in more outside-the-box-type innovations or are combining prior arts in an unique way that has not been tried by others. The Med group can be considered as a firm that is developing technologies close to the standard, which is class-level average. The Med group includes many of the big firms holding many patents. The average firm size by total assets for the Med group is $8.2 billion, whereas the Low group has an average of $2.6 billion in assets, and the High group has total assets of approximately $2 billion. The Med group has 74 patents on average while the Low and High group have approximately nine patents. However, the relative forward citations level has different implications. The High *TSL* tercile has the highest relative forward citations and the relationship is linear. Thus, patents with higher *TSL* seem to have greater technological impact than the other two groups after removing the technological class-level mean effect. Additionally, the High group expenses slightly more on R&D by 1 to 2 percentage points compared to their total assets. This is possible evidence that higher *TSL* implies high cost – a high technological return relationship. However, on average, the High group has the lowest profitability by ROA and gross profitability. Thus, it is not clear from the summary statistics if there is any linear relationship between *TSL* and operational success. Additionally, both the High group and Low group have a slightly higher market-to-book ratio than the Med group. The High group even has a slightly higher market-to-book ratio than the Low group. This could be a confounding effect of firm size, but we cannot rule out that high *TSL* is associated with growth options after controlling for the size effect. In terms of asset tangibility, the High group also has less tangible assets. Interestingly, the High group has fewer intangible assets than the other two groups while they spend more on R&D. This implies that the High group conducts more internal development than simply buying technologies from external firms. Based on the summary statistics result, the only notable difference between the High *TSL* group and the Low *TSL* group stems from different levels of *TSL*. Therefore, any pattern documented later in this paper could be plausibly considered as consequences of different *TSL*.

Table 1.5: **Summary Statistics**

The sample period is from **1976** to **2010**. A patent's *TSL* measures the average technological distance between an originating patent's class and the classes of cited patents. It is adjusted to be centered around the class's average distance. A firm's *TSL* is the average of the patent-level measure across all issued patents in a given year. It is between -1 (within-class innovation) and 1 (outside-the-box innovation). The average economic value of a patent is from Kogan et al. (2017), which is based on the three-day abnormal return around the patent issue date. The average forward citations measure how many times an originating patent receives forward citations from other patents compared to its technological class mean of forward citations. All the book values are based on fiscal-year-end values. Market cap is measured as share price times the number of shares outstanding. The book value of equity is defined as shareholder equity(SEQ) plus deferred taxes($TXDITC$, if available), minus preferred stock($PSTK$). If SEQ is not available, I use common equity(CEQ) plus preferred equity($PSTK$) or assets(AT) minus liabilities(LT). If $PSTK$ is not available, I use redemption value($PSTKRV$) or liquidation value($PSTKL$). Market value of equity for market-to-book ratio is measured at the end of the fiscal year (month t). Intangible capital is measured by total intangible assets($INTAN$) minus intangible amortization(AM). R&D expenses(XRD). ROA is measured by net income (NI) divided by total assets(AT). ROE is NI divided by common equity. The definitions of gross profit and operating leverage follow [NovyMarx, 2013], [Novy-Marx, 2011] each. Book values are winsorized at 1%.

	High		Med		Low	
VARIABLES	MEAN	SD	MEAN	SD	MEAN	SD
TSL	0.212	0.134	0.005	0.035	-0.185	0.105
Total Assets	1996.749	9685.035	8153.777	30817.300	2630.356	14182.270
Size	2095.740	9945.675	7901.860	26591.600	2917.589	14568.000
ME/BE	2.825	5.971	2.710	5.199	2.814	5.689
ROA	-0.076	0.422	-0.020	0.327	-0.054	0.410
ROE	-0.090	0.720	-0.044	0.794	-0.087	0.979
Gross Profits / Total Assets	0.343	0.316	0.352	0.272	0.357	0.302
Leverage	0.617	0.476	0.665	0.400	0.627	0.485
Operating Leverage	1.065	0.587	0.983	0.494	1.069	0.626
Sales & General Expense / Total Assets	0.349	0.298	0.292	0.233	0.338	0.328
Cost of Goods Sold / Total Assets	0.714	0.522	0.688	0.449	0.725	0.532
R&D expense / Total Assets	0.091	0.150	0.078	0.123	0.083	0.138
Property, Plant, and Equipment / Total Assets	0.251	0.177	0.270	0.174	0.253	0.178
CAPEX / Total Assets	0.059	0.055	0.061	0.049	0.059	0.053
R&D expense / PPENT	1.222	3.726	0.946	3.278	1.189	4.011
Intangible Assets / Total Assets	0.089	0.138	0.094	0.136	0.092	0.141
Average Forward Citations (Relative to Tech Class Mean)	152%		128%		126%	
Number of Patents	8.8316		74.6155		8.8364	
Number of Firmyears	14,029		13,725		13,536	

To understand the basic characteristics of *TSL* at the firm level, I choose *Immersion Corp* as an example firm. Figure 1.2 shows a *TSL* time series of *Immersion Corp*. The company was founded by Louis Rosenberg in 1993 and has been developing touch feedback technology, also known as haptic technology, used in many devices such as smartphones and gaming tools. Haptic or kinesthetic communication recreates the sense of touch by applying forces, vibrations, or motions to the user[12]. Therefore, to develop haptic technology for electronics in a broad sense, knowledge from different fields such as neuroscience, and psychology is required. *Immersion Corp* is a licensor of this technology. The firm is in the top tercile portfolio based on *TSL*. In its first public year following its IPO in 1999, the company had high *TSL*. Their technology was mainly used in gaming devices. In the following year, the TSL decreased slightly but started to increase again from year 2002 and peaked in 2007 when the first iPhone was released and many other smartphone manufacturers followed. In this example, a high *TSL* represents one of the biggest innovations in the last decade. Then, in 2013, the companies *TSL* decreased to a record low as the smartphone market matured. Here, I compare the time series pattern of the *TSL* with that of the relative forward citations to see if *TSL* is related to technological impact and applicability. The relative forward citation is computed by dividing the number of forward citations of an originating patent by the total number of forward citations of all patents in the originating patent's class. Then, I take the average across all patents in a given year. In this way, I remove the technological class-level fixed effect. As in Figure 1.2, relative forward citations closely follow the time series pattern of *TSL*. For example, patents issued in the year 2007 receive the highest relative forward citations until year 2015. Patents with higher *TSL* obtain relatively more forward citations than the patents with lower *TSL*. Thus, a higher *TSL* implies a greater technological impact if forward citation is considered a good proxy for technological impact. Moreover, *TSL* is persistent for a firm in the top tercile portfolio such as *Imersion Corp*, which weakly implies that *TSL* might be related to some systematic risk factor.

Table 1.6 shows the results from panel regression of 60-month excess return volatility on *TSL* and other controls. As stated in section 1.3, volatility is a simple risk measure widely used in practice. If the riskiness of a novel R&D strategy with high *TSL* has an effect on equity risk, it is plausible that *TSL* predicts return volatility positively. Column (1) in the table shows that *TSL*

[12] For example, when a user touches on-screen keypads on a smartphone there is feedback in the form or a weak and short vibration

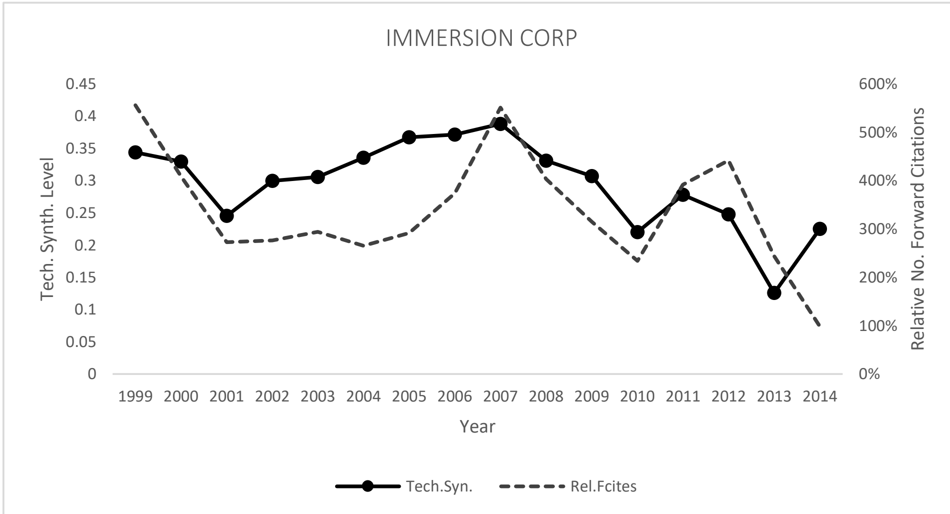

Figure 1.2: **High *TSL* example: Immersion Corporation**
Figure 1.2 shows the time series *TSL* and relative forward citations of Immersion Corporation(*NASDAQ:IMMR*). Immersion Corporation was founded by Louis Rosenberg and has been developing touch feedback technology, also known as haptic technology which is used in many devices such as smart phones. This firm is in the top decile in terms of *TSL* of all sample firms.

predicts 60-month excess return volatility positively. If a firm's *TSL* increases by 1, its 60-month volatility increases by 2 percent. Similarly, *TSL* also predicts systematic volatility positively. A one-unit increase in *TSL* leads to a 3.67 percent increase in the 60-month systematic volatility. *TSL* also positively predicts an increase in ROA volatility. Firms with highly novel technologies will experience more volatile profitability in the future. Interestingly, *TSL* has no significant association with idiosyncratic volatility. All the regressions include year fixed effects. Standard errors are clustered by firm to remove further time-series correlations in error within a firm. Although the economic impact of *TSL* on 60-month excess return volatility is not strong, the results in Table 1.6 imply that higher *TSL* could be associated with higher equity risk for a firm, particularly systematic risk. Although not reported here, *TSL* helps explain heterogeneous systematic volatility in the cross-section. Therefore, it is reasonable to state that *TSL* is associated with equity risk dynamically as well as cross-sectionally.

Table 1.7 first confirms that *TSL* has predictability in the cross-section of stock returns. The Fama–MacBeth cross-sectional regression results show that *TSL* positively predicts excess stock returns. With basic control variables such as size and market-to-book ratio, the relationship between *TSL* and stock return is only significant at the 10% level. However the significance level increases

Table 1.6: **Volatility Panel Regression**

This table shows whether *TSL* predicts 60-month excess return volatility, idiosyncratic volatility, and ROA volatility. The sample period is from **1976** to **2010**. Sixty-month excess return volatility is a forward-looking variable that is simply the log of the standard deviation of monthly returns within 60 months from the base year. Sixty-month idiosyncratic volatility is also a forward-looking variable, which is based on the Carhart 4-factor model and log values. Similarly, systematic volatility is defined as the standard deviation of predicted returns according to the 4-factor model, and it is in log values. For ROA volatility, I use the five-year standard deviation because of limited data frequency. A firm's *TSL* is the average of the patent-level measure across all issued patents in a given year. It is between -1 (within-class innovation) and 1 (outside-the-box innovation).. A firm's *TSL* is the average of the patent-level measure across all issued patents in a given year. It is between -1 (within-class innovation) and 1 (outside-the-box innovation). All the book values are based on fiscal year-end values. Size is measured using share price and the number of shares outstanding. The book value of equity is defined as shareholder equity (SEQ) plus deferred taxes($TXDITC$, if available), minus preferred stock($PSTK$). If SEQ is not available, I use common equity(CEQ) plus preferred equity($PSTK$) or assets(AT) minus liabilities(LT). If $PSTK$ is not available, I use redemption value($PSTKRV$) or liquidation value($PSTKL$). Market value of equity for book-to-market ratio is measured at the end of the fiscal year and in natural log values. ROA is defined as profit (OIBDP) over total assets (AT). Standard errors are clustered at the firm level. The detailed empirical methodology is explained in Section 1.3.3.

VARIABLES	(1) Total Vol.	(2) Idio Vol.	(3) Sys Vol.	(4) ROA Vol.
TSL	0.020*	-0.049	0.036**	0.008*
	(0.011)	(0.138)	(0.018)	(0.004)
Size	-0.051***	-0.045	-0.0400***	-0.001
	(0.011)	(0.046)	(0.013)	(0.002)
BE / ME	-0.056***	0.095	-0.059***	0.002
	(0.015)	(0.078)	(0.017)	(0.004)
Leverage	0.127***	0.861***	-0.040	0.024
	(0.043)	(0.270)	(0.064)	(0.014)
ROA	-0.161***	1.186***	-0.194***	
	(0.039)	(0.308)	(0.048)	
R&D / AT	0.040	0.280	0.033	0.131**
	(0.053)	(0.573)	(0.077)	(0.053)
Intangibles / AT	-0.023	-1.602***	0.057	0.077***
	(0.056)	(0.314)	(0.091)	(0.026)
Constant	3.936***	3.403***	3.338***	0.084***
	(0.056)	(0.263)	(0.072)	(0.014)
Observations	33,044	32,972	32,972	29,991
Adjusted R-squared	0.735	0.111	0.503	0.696
Year FE	Yes	Yes	Yes	Yes

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

with the inclusion of more proper control variables related to innovation and a firm's fundamental characteristics. A one-unit increase in *TSL* increases, from -0.5 to 0.5, is related to higher excess returns of from 3.18 percentage points to 5.1 percentage points (annualized).

Is *TSL* explained by conventional factors? To answer this question, I conduct a sorting analysis. Table 1.8 shows whether the abnormal return difference between a high *TSL* portfolio and a low *TSL* portfolio is statistically different from zero. The return of the zero-cost *TSL* portfolio is not explained by conventional 4-factors. The zero-cost *TSL* portfolio has an abnormal return of 2.532 percent (annualized) with equal-weighting and 3.12 percent (annualized) with value-weighting. This implies that *TSL* is not explained by conventional 4-factors and should be associated with different types of risk sources.

1.5.4 The channel

A channel I find is that a high *TSL* R&D strategy can affect systematic risk and stock returns in a direct channel through covariance between the stock returns and aggregate knowledge capital productivity. As seen in the previous sections, high *TSL* patents are more likely to have more significant technological impact ex-post. Therefore, they are more likely to be used by many other patents in the future. For example, [Trajtenberg et al., 1997] shows that patents with high technological distance have greater applicability in a sense that those patents are more likely to be used by technologies in a variety of fields. When the economy is strong, there will be a greater number of inventors and firms developing their technologies or products based on high *TSL* patents. On the other hand, when an economy is in recession, a significant population of the inventors and firms exit the economy. Therefore, high *TSL* patents are used less. Given these findings, high *TSL* technologies are likely to have pro-cyclical productivity. Thus, firms with high *TSL* technologies could be more subject to aggregate economic risk than firms with low *TSL* technologies.

Consistent with the above reasoning, I find patent-level evidence that patents with high *TSL* obtain more forward citations when aggregate-level innovation is more active. This shows that *TSL* could be a proxy for exposure to aggregate level innovations. Table 1.9 shows whether *TSL* is positively associated with exposure to aggregate-level innovation activities at the patent level. As shown in the Table 1.9, *TSL* positively predicts a higher level of covariance between individual

Table 1.7: **Cross-Sectional Regression**

This table shows the cross-sectional regression results using monthly returns. All returns are **excess returns**, and the unit is ***percent***. The sample period is from **1976** to **2010**. A firm's *TSL* is the average *TSL* across all issued patents in a given year. It is between -1 (within-class innovation) and 1 (outside-the-box innovation). All the book values are based on fiscal year-end values. Book values at the end of the fiscal year t are matched with monthly stock returns in year $(t + 1)$. Size is measured using share price and the number of shares outstanding. The book value of equity is defined as shareholder equity(SEQ),plus deferred taxes($TXDITC$, if available) minus preferred stock($PSTK$). If SEQ is not available, I use common equity(CEQ) plus preferred equity($PSTK$) or assets(AT) minus liabilities(LT). If $PSTK$ is not available, I use redemption value($PSTKRV$) or liquidation value($PSTKL$). Market value of equity for market-to-book ratio is measured at the end of the fiscal year. R&D expenses(XRD) are used as a proxy for intangible investment and are normalized by total assets. I exclude no R&D expense firm years from the sample. Industry fixed effect is controlled by dummy variables based on the three-digit sic code.

VARIABLES	(1) Excess Return	(2) Excess Return	(3) Excess Return	(4) Excess Return
TSL	0.265*	0.369**	0.425**	0.366**
	(0.160)	(0.179)	(0.187)	(0.179)
Knowledge Capital / AT		0.706**	0.592**	0.699**
		(0.288)	(0.271)	(0.289)
PPENT / AT		-0.148	-0.0170	-0.229
		(0.282)	(0.293)	(0.281)
β_{MKT}			0.0882	
			(0.121)	
β_{HML}			0.108	
			(0.0721)	
β_{SMB}			0.104	
			(0.0720)	
β_{UMD}			-0.100	
			(0.0785)	
R&D /AT		3.290**	1.577	3.222**
		(1.278)	(1.165)	(1.289)
Leverage		0.921***	0.195	0.845***
		(0.278)	(0.242)	(0.275)
ROA		3.233***	2.081***	3.289***
		(0.431)	(0.458)	(0.434)
Size	-0.124**	-0.177***		-0.164***
	(0.0507)	(0.0433)		(0.0453)
BE / ME	0.467***	0.586***		0.599***
	(0.0788)	(0.0846)		(0.0855)
Momentum	0.565***	0.383**		0.379**
	(0.180)	(0.172)		(0.171)
Operating Leverage				0.138
				(0.120)
Constant	1.193	2.721*	-0.767	2.826*
	(0.966)	(1.510)	(1.038)	(1.458)
No. of Firms	4,627	3,549	3,441	3,548
R-squared	0.201	0.217	0.219	0.219
Industry FE	YES	YES	YES	YES

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 1.8: **Sorting Analysis: *TSL* Sorted Portfolios**

This table contains the time series regression of *TSL*sorted portfolio returns. The sample period is from 1976 to 2010. A firm's *TSL* is the average of the patent-level measure across all issued patents in a given year. It is between -1 (within-class innovation) and 1 (outside-the-box innovation). Other factor variables are obtained from Kenneth French's website.

Variable	Low	(1)	(2)	(3)	High	H-L
			Panel A: Equal-weighted portfolios			
α	0.238**	0.331	0.369***	0.339***	0.448***	0.211**
	(0.096)	(0.094)	(0.084)	(0.098)	(0.109)	(0.095)
MKTRF	1.062***	1.084***	1.080***	1.111***	1.052***	-0.010
	(0.022)	(0.021)	(0.019)	(0.022)	(0.025)	(0.022)
HML	-0.050	-0.012	-0.028	-0.017	-0.103***	-0.053
	(0.034)	(0.034)	(0.030)	(0.035)	(0.039)	(0.034)
SMB	0.934***	0.820***	0.615***	0.824***	1.019	0.085***
	(0.032)	(0.031)	(0.027)	(0.032)	(0.036)	(0.031)
UMD	-0.170***	-0.220***	-0.220**	-0.228***	-0.244***	-0.074***
	(0.021)	(0.020)	(0.018)	(0.021)	(0.024)	(0.021)
R-squared	0.918	0.918	0.927	0.916	0.904	0.0415
			Panel B: Value-weighted portfolios			
α	0.787***	0.715***	0.692***	0.740***	1.047	0.260*
	(0.111)	(0.102)	(0.076)	(0.094)	(0.122)	(0.158)
MKTRF	0.924***	0.997***	0.962***	0.956***	0.992	0.068*
	(0.025)	(0.023)	(0.017)	(0.021)	(0.028)	(0.036)
HML	-0.177***	-0.159***	-0.178***	-0.072***	-0.220***	-0.042
	(0.040)	(0.036)	(0.027)	(0.033)	(0.044)	(0.056)
SMB	0.231***	-0.177***	-0.161***	0.095***	0.130***	-0.101*
	(0.037)	(0.033)	(0.025)	(0.031)	(0.040)	(0.052)
UMD	0.085***	0.026	-0.039***	-0.090**	-0.030	-0.115***
	(0.024)	(0.022)	(0.017)	(0.020)	(0.027)	(0.034)
R-squared	0.819	0.840	0.901	0.860	0.807	0.0386

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

patent productivity and aggregate-level productivity. Even after controlling for some alternative measures regarding the qualitative characteristics of innovation, *TSL* still maintains a positive and significant relation with β_{fcites}. Altogether, high *TSL* patents become more technologically useful (productive) when overall innovation activity measured by the amount of forward citations is high. This finding could be expanded to the firm-level relationship between *TSL* and aggregate productivity of knowledge capital. Firms with high *TSL* are more subject to exogenous shock to the aggregate productivity of knowledge capital, and they have higher stock returns because of this co-movement.

Table 1.10 shows the correlation coefficient between aggregate consumption shock and knowledge capital productivity shock. There is a correlation coefficient of 0.309, which is statistically significantly different from zero. This positive correlation implies that knowledge capital productivity could also be pro-cyclical.

If the high *TSL* of a firm proxies for sensitivity to the knowledge capital productivity risk as we see from patent-level evidence in Table 1.9, the firm should have higher stock returns as a consequence of pro-cyclicality of knowledge capital productivity. However, this should be interpreted carefully because there could have an accumulation of non-negligible estimation errors in each stage. Additionally, this is an indirect way of showing if a risk factor carries a positive price of risk.

To resolve this concern, Table 1.11 adds more direct evidence that knowledge capital productivity carries a positive price of risk. I use 130 test assets including 25 size-BM portfolios, 25 BM-investment portfolios, 25 size-investment portfolios, 25 size-profitability portfolios, 25 profitability-investment portfolios, and 5 TSL portfolios. I first estimate betas by running time-series regression and then running cross-sectional regression to determine if the slope coefficient of the knowledge capital productivity shock is positive and significant. Consistent with the results in Table 1.10, the slope coefficient for the knowledge capital productivity shock is positive and statistically significant. Although the magnitude should be interpreted carefully due to possible measurement errors, the estimated price of risk for knowledge productivity shock is 0.636 percent per month (7.632 percent annualized), which is economically significant. This directly implies that firms positively exposed to knowledge capital productivity shock should have higher stock returns. Also, together with the finding in Table 1.10 that the knowledge capital productivity shock has positive correlation at 11.6

Table 1.9: **Systematic Risk: Patent-level Evidence**

This table contains the regression of patent citations beta (β_{fcites}) on *TSL* and other control variables. The sample period is from 1976 to 2014. *TSL* measures the average technological distance between an originating patent's class and the class of cited patents. A firm's *TSL* is the average of the patent-level measure across all issued patents in a given year. It is between -1 (within-the-class innovation) and 1 (outside-the-class innovation). Economic value is estimated from the stock market reaction to patent issuance as in [Kogan et al., 2017], and the data is obtained from Noah Stoffman's website. The number of backward citations is the total number of backward references of an originating patent. The number of self-citations is the number of backward citations made by the same assignee. The number of external citations is the number of backward citations made by examiners. The Trajtenberg distance measures a similar quality as *TSL*, but it uses the simplified distance between the originating patent's class and its backward citations([Trajtenberg et al., 1997]). Please see section 1.7 for a detailed explanation. The generality score is the number of different technological classes of the patents citing the originating patent. Innovative Originality at the patent level is the number of unique technological classes of patents referred to by an originating patent up to secondary classes following [Hirshleifer et al., 2018]. The detailed empirical methodology is explained in Section 1.3.4.

VARIABLES	(1) β_{fcites}	(2) β_{fcites}
TSL	0.003***	0.005*
	(0.001)	(0.003)
Economic value	3.94e-04***	1.89e-04***
	(1.16e-04)	(5.80e-05)
No. Backward Citations	0.001***	0.001***
	(1.62e-04)	(1.56e-04)
No. Self Citations	-0.027***	-0.003
	(0.009)	(0.004)
No. External Citations	0.028*	0.046***
	(0.016)	(0.010)
Trajtenberg distance		-0.059***
		(0.010)
Generality score		0.045***
		(0.002)
Innov. Originality		-0.004***
		(4.05e-04)
Constant	-0.005*	-0.146***
	(0.003)	(0.006)
Observations	1,026,802	1,026,802
R-squared	0.058	0.386

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 1.10: **Knowledge Capital Productivity Shock Correlation**

This table shows the correlation coefficients of knowledge capital productivity shock with aggregate consumption shock and zero-cost *TSL* portfolio returns. Knowledge capital productivity (or factor share) is obtained from year-by-year cross-sectional productivity regression. Firm level output is measured by net sales (SALE) deflated in 2009 dollars. Input factors are physical capital, knowledge capital, other intangible capital, and labor. Since compustat Labor(EMP) is the total number of employees, I divide output and inputs by the number of employees and run per-capita productivity regression. Firm level intangible capital is obtained from [Peters and Taylor, 2016]. For physical capital stock, I use COMPUSTAT PPEGT. The detailed empirical methodology is explained in Section 1.3.4.

	Knowledge capital productivity shock
Consumption shock	0.309**
	(0.0116)
High TSL minus Low TSL	0.116**
	(0.0170)

P values in parentheses

percent with the returns of High-minus-Low-TSL portfolio, it is reasonable to conjecture that the TSL's explanatory power will be partially subsumed by this knowledge capital productivity risk variable.

Table 1.12 repeats the Fama–MacBeth regression in Table 1.7 adding the knowledge capital productivity beta(β_{know}). If the primary source of the return difference between high *TSL* stocks and low *TSL* stocks is co-movement with aggregate knowledge capital productivity, adding β_{know} in the Fama–MacBeth regression should reduce the statistical power of *TSL* dramatically because β_{know} is a direct measure of co-movement between a stock return and aggregate knowledge capital productivity. As in Table 1.12, when including β_{know} in the Fama–MacBeth regression, the coefficient of *TSL* decreases from 0.369 to 0.298 (by 19 percent), and so does its statistical power. On the other hand, the β_{know} is statistically significant at 10 percent level, and the marginal effect of it is around 0.84 percent annualized. So it seems that the part of TSL's explnatory power which is correlated with knowledge capital productivity risk is subsumed by the β_{know}.

In addition, Figure 1.3 shows that *TSL* portfolios are aligned closely to the 45 degree line, which implies knowledge capital productivity risk helps explain returns of *TSL* portfolios. Compared to the graph in the left of Figure 1.3, the right panel, which use four factors plus knowledge capital factor, clearly fits test assets better. Especially, the fitness of *TSL* portfolios are highly improved than the four factor model. This also supports that knowledge capital productivity risk explains

Table 1.11: **Fama–MacBeth Regression**

This table contains the cross-sectional regression of test assets on beta to knowledge capital productivity shock and other conventional betas. The sample period is from 1976 to 2010. β_{know} is measured by regressing excess returns on knowledge capital productivity shock. For the definition of knowledge capital productivity shock, please see Table 1.10. The detailed empirical methodology is explained in Section 1.3.4.

VARIABLES	(1) Excess Ret
β_{know}	0.762***
	(0.165)
β_{mkt}	0.421**
	(0.128)
β_{smb}	0.389***
	(0.048)
β_{hml}	0.441***
	(0.060)
β_{umd}	2.416***
	(0.361)
Constant	0.381
	(0.111)
Observations	130
R-squared	0.66

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 1.12: **Fama–MacBeth Regression**

This table contains the Fama–MacBeth regression of individual stock returns on *TSL* and other control variables. The sample period is from 1976 to 2010. β_{know} is measured by regressing excess returns on knowledge capital productivity shock. For the definition of knowledge capital productivity shock, please see Table 1.10. For the definitions of other control variables, please refer to Table 1.7. This column corresponds to column (2) of Table 1.7.

VARIABLES	(1) Excess Ret
TSL	0.298
	(0.193)
β_{know}	0.070*
	(0.039)
Knowledge Capital / AT	0.610**
	(0.300)
PPENT / AT	0.054
	(0.291)
Size	-0.142***
	(0.046)
BE / ME	0.100
	(0.094)
Momentum	-0.367*
	(0.200)
R&D / AT	1.893
	(1.375)
Book Leverage	0.605**
	(0.290)
ROA	2.731***
	(0.413)
Observations	3,448
R-squared	0.259
Industry FE	YES

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

the difference between stock returns of High TSL and Low TSL portfolios. Although not reported, the difference between β_{know} of highest quintile TSL portfolio (0.133) and β_{know} of lowest quintile TSL portfolio (-0.048) is 0.18, and it implies around 1.65 percent of returns when multiplied by the estimated price of risk in Table 1.11. This is about 65 percent of the abnormal return of equal weighted portfolio, which is economically significant.

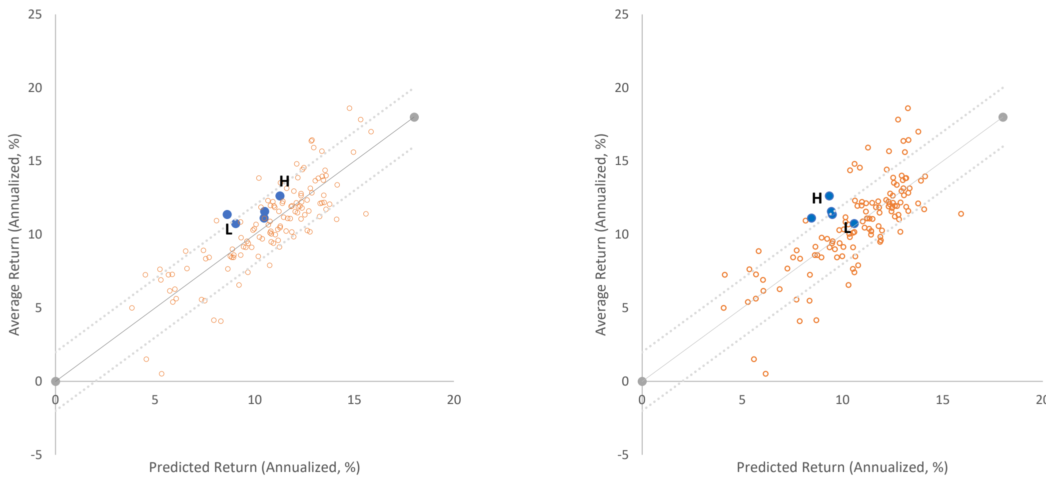

Figure 1.3: **Fitted Returns versus Average Portfolio Returns**
Figure 1.3 shows average returns of test asset portfolios versus fitted returns of the portfolios. (A) The conventional 4 factors and knowledge capital productivity factor in 1.11 are used to fit the average portfolio returns.(B) The conventional 4 factors are used to fit the average portfolio returns. L denotes lowest TSL portfolio and H denotes highest TSL portfolio. Orange bubbles (empty circles) represent test asset portfolios other than TSL portfolios. TSL portfolios are represented by blue dots (filled circles). Gray straight line is 45 degree line.

1.5.5 Alternative Explanations

The evidence in the previous sections implies that TSL could be associated with systematic risk rather than idiosyncratic risk. Also, I show that knowledge capital productivity risk could be a source of systematic risk with various supporting evidence. However, I still should not rule out other possible explanations for the findings.

As briefly mentioned in the introduction, one possible explanation is the theory suggested in [Berk et al., 2004]. I first find that TSL is positively associated with extreme outcomes, which means it is a proxy for technological risk (Table 1.1, 1.2, 1.3, and 1.4). Because patent issuance is one of the stages in an R&D project, the technological risk proxied by TSL representing an ex-ante characteristic of issued patents is related to failure risk in the development stage. According to

[Berk et al., 2004], this type of failure risk can indirectly affect the risk premium of a firm, which is a collection of multistage R&D projects, by increasing the threshold for the option to invest in the next stage of the project. Consistent with the theory's prediction, I find that high *TSL* is associated with high stock returns and systematic volatility. Therefore, pursuing novel R&D strategy indirectly increases stock returns.

Alternatively, the positive association between *TSL* and systematic risk could stem from financial constraint risk. High *TSL* projects may be hard to finance with debt because they require massive investments due to the technological difficulty, so firms with high *TSL* are likely to be more subject to financial constraint risk. According to [Whited and Wu, 2006], the stock returns of financially constrained firms move together and earn higher returns. For example, in Table 1.4, high *TSL* is positively related to high funding amount, which implies that high *TSL* projects are expensive in the first place. Also, this could be a reason why the startups with high *TSL* do more IPOs (Table 1.3) in terms of that they have to go to public equity market to raise funds because they are no longer able to finance more with debt financing.

Another possible explanation is a behavioral explanation. Similar to [Hirshleifer et al., 2018], high *TSL* projects could be more difficult to evaluate. By its definition, to draw new knowledge from previous knowledge in the very different field is probably something that has not been tried many times in the past. Therefore investors might lack enough information to assess the values of high *TSL* projects. In this case, stocks with high *TSL* could be discounted and have higher stock returns. However, as in Table 1.1, *TSL* does predict the higher probability of winning innovation awards, which implies the existence of people who can assess the fair value of high *TSL* projects in the market. It could be contradicting evidence to the valuation uncertainty hypothesis at least with the sample firms in this study.

1.6 Robustness

1.6.1 Correlation with other measures

Table 1.14 repeats the R&D100 awards regression in Table 1.1 replacing *TSL* with alternative measures such as Trajtenberg distance and Innovative Originality. [Trajtenberg et al., 1997] suggests a backward-looking measure that shows if a patent combines remote technological areas. If the originating patent has the same three-digit technological class with a backward-cited patent, the weight is 0. If the originating patent has the same two-digit technological class but a different three-digit class with a backward citation, 0.33 of the weight is given. If the first digit of the originating patent's class and the first digit of the referred patent's class are the same, the weight is 0.66. If the first digit of the originating patent's class and the first digit of the referred patent's class are different, the weight is 1. Another alternative measure is Innovative Originality, which is the total number of unique technological classes up to secondary classes.

In Table 1.13, *TSL* has a correlation coefficient of 0.706 with the technological distance computed as in [Trajtenberg et al., 1997]. This shows that both capture similar qualitative aspects but they are still differentiated. This is because *TSL* uses a more accurate way of measuring technological distance by using the historical citations network rather than simply and arbitrarily assigning distances based on similarity in classification codes. The second row shows how much *TSL* is correlated with the originality measure. As [Trajtenberg et al., 1997] distinguishes originality and technological distance, *TSL* is also distinguishable from Innovative Originality suggested by [Hirshleifer et al., 2018] because they have only 0.232 of the correlation coefficient. The originality measures how many various ingredients are used while *TSL* measures how different the ingredients are from those that are newly created. Therefore, *TSL* measures the different dimensions of technology not captured by Trajtenberg distance or Innovative Originality.

Table 1.13: **Correlation Table**

This is a correlation table of *TSL* and other patent characteristics, technological distance in [Trajtenberg et al., 1997], and Innovative Originality in [Hirshleifer et al., 2018]

	TSL
Trajtenberg distance	0.706
Innovative Originality	0.232

Table 1.14 shows whether these two alternative measures capture what *TSL* proxies, which is technological breakthrough. In panel(1) of Table 1.14, patent-level Innovative Originality does not show statistical predictive power for the likelihood of earning an R&D 100 award. In panel(2), the Trajtenberg distance predicts award-winning probability, but the statistical power is not as strong

as that of *TSL*. These results imply that *TSL* could be a better proxy for technological breakthrough than the alternative measures.

Table 1.14: **Alternative Measures of *TSL*: Breakthrough Regression**
This table repeats the R&D100 awards regression in Table 1.1 replacing *TSL* with alternative measures such as the Trajtenberg distance and Innovative Originality. The sample period is from 1976 to 2014. The Trajtenberg distance measures similar qualities as *TSL* but with a simpler weighting scheme. Innovative Originality at the patent level is the total number of unique technological classes of all backward citations of the originating patent up to secondary classes.

VARIABLES	(1) R&D 100 Awards(Title)	(2) R&D 100 Awards(Title)
Trajtenberg Distance		0.097*
		(0.058)
Innovative Originality	0.003	
	(0.004)	
Num. of Backward Citations	5.77e-04	9.93e-04***
	(9.42e-04)	(3.75e-04)
Self Citations	0.164	0.169
	(0.113)	(0.112)
External Citations	0.355	0.365
	(0.596)	(0.596)
Constant	-2.671***	-2.694***
	(0.109)	(0.109)
Observations	146,624	146,624
Year FE, class cluster	Yes	Yes

Robust standard errors in parentheses
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.15 shows cross-sectional regression results when including ohter major innovation related variables such as Innovative Efficiency and Innovative Originality. Innovative Efficiency measures a quantitative characteristic of innovation by total number of output patents. As described above, Innovative Originality measures a qualitative characteristic of innovation which focuses on how broad the ingredient technology fields are. As shown in Table 1.15, even after controlling for other innovation characteristic, *TSL* still can positively predict stock returns at 10 percent significance level. This evidence supports that *TSL* is distinguished from (1) other quantitative characteristic and (2) other qualitative characteristic.

Table 1.15: **Alternative Measures of *TSL*: Fama-Macbeth Regression**

This table repeats the Fama-Macbeth Cross-Sectional regression adding other innovation related variables such as Innovative Originality and Innovative Efficiency([Hirshleifer et al., 2013, Hirshleifer et al., 2018]). The sample period is from 1976 to 2014. Innovative Originality at the patent level is the total number of unique technological classes of all backward citations of the originating patent up to secondary classes. Innovative Efficiency is natural log of one plus total number of patents issued stnadardized by past R&D investments.

	(1)
VARIABLES	Excess Return
TSL	0.319*
	(0.176)
Inn. Org	0.0363
	(0.0663)
IE	0.0780
	(0.140)
Knowledge Capital/ AT	0.665**
	(0.300)
R&D / AT	3.820***
	(1.297)
PPENT / AT	-0.372
	(0.266)
Operating Lev	0.133
	(0.115)
Size	-0.158***
	(0.0430)
BE/ME	0.626***
	(0.0809)
Momentum	0.377**
	(0.165)
Leverage	0.884***
	(0.255)
ROA	3.172***
	(0.405)
No. Firms	3,258
R-squared	0.216

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

1.7 Conclusions

Despite a long history of literature discovering the relationship between innovation and the cross-section of stock returns, there are few academic works that focus on how the qualitative aspects of innovation affect firm risk and stock returns. This paper contributes to the literature by shedding light on a qualitative characteristic of innovation represented by technological distance. Additionally, I add evidence that the new pattern could be related to systematic risk.

First, this paper contributes to the innovation literature by suggesting a new novelty measure of innovation, which shows the extent to which a new patented technology is built upon extremely different prior technologies and discovers its association with more risky outcomes. I provide patent and startup-level evidence using novel data obtained from Crunchbase.

Second, this paper contributes to the asset pricing literature by (1) showing a new cross-sectional pattern related to an ex-ante qualitative characteristic of innovation and (2) providing evidence that the pattern is related to systematic risk.

For firm level analyses, I find that higher *TSL* is associated with higher stock returns in the cross-section. I also show that high *TSL* is positively related to high future return volatility (total and systematic) but find no relationship with idiosyncratic volatility. This implies the possibility that *TSL* is priced in the cross-section of stock returns and subject to some type of systematic risk source. At least, the finding supports meaningful association between *TSL* and systematic equity risk.

Next, I find evidence that high *TSL* patents are more productive when the aggregate innovation level is high. In addition, knowledge capital productivity risk has a positive slope coefficient in the cross-sectional regression with various tested asset portfolios. Finally, when the firm-level beta to knowledge capital productivity shock is included in the Fama–MacBeth regression, *TSL*'s explanatory power becomes weak, which means the effect of *TSL* is subsumed by knowledge capital productivity risk.

The empirical evidence supports a channel that explains positive association between *TSL* and systematic equity risk. Firms with high *TSL* can have high co-movement with aggregate innovation productivity because novel innovation with high *TSL* is related to significant technological impact in the sense that it is referred to by more post-technologies. Coupled with the finding that innovation productivity risk has a positive price of risk, firms with high *TSL* should have higher stock returns because of positive covariance with aggregate innovation productivity.

Alternatively the evidence also supports that the failure risk associated with novel R&D strategy can indirectly change the risk premium([Berk et al., 2004]). The finding that high *TSL* is associated with a higher likelihood of patent failure and high stock returns supports this theory. Also this paper still has a room to explore in terms of distinguishing between the knowledge productivity risk channel and the alternative channels briefly discussed in the Section 1.5.5.

Overall, my results show a possibility that a qualitative dimension of innovation, the technological distance (*TSL*), is aggregated at firm-level and priced in the cross-section of stock returns through systematic risk related to knowledge capital productivity.

Chapter 2

Intangibles and Equity Risk Collateral Channel

2.1 Introduction

The role of intangible capital and investments has been increased in the economy. [Corrado and Hulten, 2014] documents that the investments in intangible capital started to exceed the investments in tangible capital. In response to these economic changes, we observe more cases that intangible capital is used as collateral for corporate loans. The history of public firms' intangible collateral usage only goes back to mid 90's, thus it can be considered as fairly recent and new phenomenon. Although the finance and economics studies have mostly assumed that intangible capital is non-pledgeable portion of a firm's assets, there are some recent studies trying to incorporate the intangible capital in the model or find empirical evidence shedding light on the important role of intangible capital and investments .([Li et al., 2014],[Peters and Taylor, 2016], [Ai et al., 2012], [Belo et al., 2014])

Also there are several studies focusing on the cases that specific intangible assets used as collateral. For example, [Mann, 2018] studies the patent collateral of US public firms and add empirical evidence that it increases debt financing. [Lumioti, 2012] also investigate how the usage of intangible collateral is related to credit risk using Dealscan data.

However, not yet many studies have focused on how the usage of intangible collateral is related to equity returns and firm risks. Recently, one possible mechanism of how the type of collateral affects stock returns and firm risk that has been studied by [Ai et al., 2018]. They show that a firm's asset composition matters for its expected stock return because collateralizability of the assets can provide insurance to economic downturn and reduce covariance of equity returns and business cycle. They first suggest a production economy model that allows two different types of assets based on collateralizability. Their model predicts that the firms with more collateralizable assets have lower stock returns than the other firms with less collateralizable assets. When a negative aggregate shock arrives, the return on pledgeable capital is less responsive to the shock than the return on non-pledgeable (intangible) capital, because pledgeable capital can relax financial constraint which binds more in this situation. From the model, they develop an empirical measure for asset collateralizability assuming that intangible assets are not be able used as collateral. They find empirical evidence consistent with their model that firms with high collateralizability have lower stock returns than the other firms with lower collateralizability. What they argue is that insurance channel dominates the leverage effect of highly collateralizable assets.

Although this collateralizability premium sounds interesting and plausible to explain the higher stock returns for firms with higher level of intangible capital (low collateralizability), one can easily find empirical evidence that their assumption that intangible assets are not pledgeable can be falsified. There are cases that the intangibles are pledged as mentioned at the beginning of this paper. Also [Mann, 2018] finds that about 38 percent of US patenting firms had pledged patents by 2013. Therefore collateralizability premium estimated in [Ai et al., 2018] could be confounded with other effects related to intangible capital's nature.

However, the fact that there is empirical evidence against the assumption in [Ai et al., 2018] does not mean that the collateralizability premium does not exist at all. In order to test whether the collateralizability premium exists, one can first draw a hypothesis regarding firms pledging intangible capital as collateral by using the collateralizability premium story. At this moment, one can slightly put aside the validity of the assumption about intangible capital's collateralizability and test whether the hypothesis based on the collatrealizability premium works in the empirical world. Therefore, this paper tries to look at the collateralizability premium in another perspective that is not tested in the [Ai et al., 2018].

Let us assume that intangible capital is not fundamentally different from tangible capital when it is pledged as collateral. Then the role of intangible capital is simply to relax financial constraint and give a firm an ability to increase borrowing. For example, [Mann, 2018] finds the empirical evidence that patent collateral increases debt financing. In other words, not only tangible collateral but also intangible collateral can provide insurance for tighter financial constraints during an economic recession. Although it will be described in the next sections, the intangible capital was pledged together with tangible capital in most of the cases, which implies that firms try to relax financial constraint more by using the intangibles in addition to the tangibles. Therefore it makes sense that intangible capital gives additional slack in financial constraints. In this case, firms who can pledge intangible capital should have lower stock returns if the collateralizability premium exists, compared to the similar firms who can not pledge intangible capital. In short, if intangible capital does not have any stark differences with tangible capital as collateral, then firms that have can pledge intangible collateral should be considered as firms with more collateralizable assets. On the other hand, if intangible collateral is something clearly distinguished from tangible collateral by its fundamental nature, it may have different effects on firm variables and risks than simply relaxing financial constraint as tangible collateral does. Thus I can test whether the collateralizability premium works for the specific set of sample firms by looking at syndicated loans backed by intangible collateral with proper matching exercise.

I first compare 'firm-months with intangible collateral' and 'firm-months without intangible collateral.' Based on the collateral type information of the LPC Dealscan database, firm-months are assigned a value of one if any loan pledging intangibles as collateral is not expired. Otherwise, firm-months are assigned a value of zero.

As for preliminary examination, I run Fama-Macbeth regression and find that intangible collateral positively predicts excess return going forward. Also, I compose a portfolio that takes long positions in intangible-collateral stocks and short positions in tangible-collateral stocks monthly. The long-short portfolio has a positive abnormal return around 8 to 10 percent (annualized) after controlling for conventional factors.

Together with the results, I also find that the firms with intangible collateral achieve similar or slightly higher level of leverage compared to the other firms with only tangible collateral. This shows that the intangibles also play a role to relax financial constraints, which is assumed impossible in the [Ai et al., 2018].

These results give an impression that the collateralizability premium might not exist or be fairly small even if it exists for the sample firms with intangible collateral. The sample firms can clearly relax financial constraint with intangible collateral, but they still earn higher abnormal returns. This is against the mechanism of the collateralizability premium.

However, in this analysis, the sample size of treated firms is small while the sample size of control firms is large. Thus, there could be some fundamental differences in the characteristics of treated firms and those of control firms. In order to resolve the problem and test the hypothesis with a more rigorous setting, I match 'firms with syndicated loans that are backed by intangible collateral' with 'firms without syndicated loans' based on several firm-level dimensions including the size of intangible capital. The key is that firms are matched based on the intangible capital level. Again, if the collateralizability premium exists, firms with intangible collateral should have lower stock returns, compared to the matched firms that do not have any syndicated loans even with similar level of intangible capital, because they can relax their financial constraint by pledging the intangibles.

I use nearest-neighbor matching method allowing one match per observation. The firms with intangible collateral are considered as treated, and they are matched with firms without syndicated loan and intangible collateral by asset size, intangible capital level standardized by total assets, R&D investment intensity, degree of financial constraint measured by HP Index, net debt over sales, and access to the public debt market.

The matching analysis shows that firms with syndicated loans backed by intangible collateral have higher stock returns than the other matched firms with similar level of intangibles but no syndicated loans. The results are robust over several subsample periods. This evidence could be against collateralizability premium. In other words, treated firms and matched firms have similar level of intangible capital, intangible investment, and financial constraint, but the treated firm can loosen the financial constraint by intangible collateral while matched firms cannot. Therefore it is expected to see lower stock returns for treated firm and higher returns for matched firm, but the evidence is the opposite.

In conclusion, the empirical evidence in this study does not support the fairly big collateralizability premium (8 percent annualized). The collateralizability premium might be confounded by the effect of intangible capital.

Lastly, I add some suggestive evidence for why the firms with intangible collateral have higher stock returns. Of course, the primary purpose of this study is to test collateralizability premium hypothesis with different empirical data that is not used by [Ai et al., 2018] and add some counterevidence to the hypothesis. What I do in this part of the analysis is to add very brief and small evidence that might let us know the direction of future studies regarding the intangible collateral.

I calculated the sensitivity of each treated(with intangible collateral) and matched(without intangible collateral) portfolio to technology adoption shock in [Lin et al., 2018] as well as knowledge capital productivity shock suggested in Chapter 1 to see whether the treated portfolio is somehow more exposed to these risks. Hypothetically, firms that can pledge intangibles as collateral could have better quality of intangibles and also higher intangible productivity and be more exposed to aggregate technology adoption shock, and that is the reason the firms have higher stock returns.

I follow [Lin et al., 2018] to estimate technology adoption shock empirically. The technology adoption shock carries a positive price of risk; therefore firms that are positively exposed to the technology adoption shock should have higher stock returns. The technology adoption shock is calculated by the new introduction of technology standards in the general technology field. Although [Lin et al., 2018] estimate the technology adoption shock at a quarterly frequency, due to data restriction, I estimate the shock at an annual frequency.

Consistent with the prediction, the portfolio of firms with intangible collateral have higher betas to intangible capital productivity and technology adoption shock. Therefore, it is possible that firms with intangible collateral have higher stock returns due to their higher exposure to aggregate technology risk.

The paper is organized as follows. Section 2.2 presents a literature review. Section 2.3 explaines the empricial methodology used in the analyses in detail. Section 2.4 describes the data used in the paper. Section 2.5 presents the findings with a discussion, and Section 2.6 concludes.

2.2 Literature Review

This study contributes to the literature about intangible collateral and how it affects firms. For example, as briefly mentioned in the previous section, [Mann, 2018] collects sample firms with patent collateral. He shows that the pledgeability of patents is positively related to increase in debt financing and innovation investment when creditor rights to the patent collateral is strongly protected.

[Hochberg et al., 2018] shows that redeployability of patent collateral matters for venture lending.

They use Dot-com crash as an unexpected capital supply shock to VCs. [Mann, 2018] also finds similar pattern that highly redeployable patents are more likely to be pledged as collateral.

Also, [Lumioti, 2012] show that intangible collateral is credit market innovation and can alleviate financial constraint using syndicated loan sample from Dealscan database. This paper focuses on the determinants of usage of intangible collateral and performance of the syndicated loans backed by intangible collateral.

While [Mann, 2018] and [Hochberg et al., 2018] focus on a specific type of intangible collateral, this study includes all other types of intangible collateral in addition to patent collateral. [Lumioti, 2012] utilizes very similar data with data in this paper, but she does not explore the relationship between intangible collateral and equity returns. The other two studies also examine questions regarding credit risk and debt financing. To my knowledge, this is the first paper investigating the relationship between intangible collateral and stock returns.

Of course, several papers have studied how incorporating intangible assets in the model changes conclusions about firm risk and investments, but they barely focus on the effect of intangible collateral.

For example, [Li et al., 2014] uses structural estimation of the q-theory augmented with intangible capital investment and show that the model captures value premium and the positive relationship between R&D intensity and stock returns better. They assume different adjustment costs and intangible-investment-specific technological change in the model.

[Peters and Taylor, 2016] also point out that the most neoclassical theory is tested only with physical capital investment and show that the classic q-theory performs better when intangible capital is included in the estimation. They propose a new Tobin's q proxy that includes intangible capital and show that it is superior than previous Tobin's q proxy.

[Belo et al., 2014] focuses on a specific type of intangible capital, which is brand capital. They develop an investment-based model that incorporates brand capital and show that firms with higher brand capital intensity have higher stock returns than firms with low brand capital intensity. The model can explain the empirical relationship between advertising investment and stock returns.

As shown in these studies, incorporation of intangible capital improves the models' explanatory power for firm investments and specific empirical patterns regarding stock returns. This paper also contributes to this literature by introducing an empirical relationship between the intangible capital and stock returns.

Lastly, this paper is also related to the literature about financial constraint risk and stock returns. The empirical relationship between financial constraint risk and stock returns documented long ago. [Lamont et al., 2015] shows that financially constrained firms' stock returns co-move, which suggest there is a common risk factor that those financially constrained firms are exposed together. [Whited and Wu, 2006] also finds co-movement between financially constrained stocks and show that the constrained firms earn higher stock returns. [LIVDAN et al., 2009] also study the effect of financial constraint risk by solving an investment-based asset pricing model with collateral constraints on debt capacity. In their model, collateral constraint provides inflexibility of capital adjustment and show financially constrained firms are riskier than unconstrained firms. [HAHN and LEE, 2009] empirically explore the similar question of how debt capacity is related to stock returns across financially constrained and unconstrained firms. They find that debt capacity matters specifically for financially constrained firms' stock returns. My study also studies whether financial constraints are related to cross-section of stock returns by testing the collateralizability premium hypothesis, which is based on the insurance effect for tighter financial constraints.

2.3 Method

First, I identify if a firm has any loans that are backed by intangible collateral by analyzing security type text information from Dealscan database. The procedure will be explained in detail in the next section (Section 2.4). However, the security type is only available at loan-facility level, so I assume that a firm is exposed to the effect of intangible collateral during the entire maturity of a loan backed by intangible collateral. For example, if a firm enters into a contract of 5-year syndicated loan and pledges their patents as collateral from April 2019 to April 2024, then the entire 60 months during the period are considered as firm-months exposed to the effect of intangible collateral, either its insurance effect or increased intangible related risk. More specifically, I assume the firm is exposed to the effect of intangible collateral during the entire fiscal year if the firm enters into a contract at any time of the year. Based on the assumption, I define an indicator variable $IP_{i,t}$ of a firm at time t as below. Going back to the example right above, the firm's $IP_{i,t}$ variable should have a value of

Table 2.1: **Comparison Groups based on Syndicated loans and Intangible collateral**

	Intangible Collateral	
Syndicated loans	(1)Yes, Yes	(2)Yes, No
	(3)No, Yes	(4)No, No

1 from the beginning of the fiscal year 2019 to the end of the fiscal year 2024. Otherwise the $IP_{i,t}$ should be 0.

$$IP_{i,t} = \begin{cases} 1 & \text{if intangible collateral is pledged in month t} \\ 0 & \text{else} \end{cases}$$

As a motivating analysis, I first compare firms pledging intangible collateral(cell (1) of Table 2.1) and firms pledging only tangible collateral(cell (2) of Table 2.1). As briefly described in the Section 2.1, intangible capital is pledged together with tangible collateral in most cases. Thus, this comparison should provide implications regarding the effects of intangible capital when pledged in addition to tangible capital.

I first run Fama-Macbeth cross-sectional regression to see whether the usage of intangible capital as collateral has predictive power on future stock returns. Since it is highly likely that financial constraint related risk can confound the result, I control for not only well-known firm-level stock return predictors but also variables related to financial constraint including credit rating. Also, loan characteristics are controlled, since the sample cross-section is fairly small due to the nature of the lower frequency of intangible collateral usage. If (1) intangible capital is not fundamentally different from tangible capital when used as collateral, and (2) collateralizability premium appropriately explains the cross-section of sample firms, it is expected for the intangible collateral variable to negatively predict excess return.

$$r_{i,t+1} = \beta_{0,t} + \beta_{1,t} IP_{i,t} + Controls_{i,t} + \epsilon_{i,t} \tag{2.1}$$

Next, I construct a long-short portfolio by taking long positions in firms with intangible collateral and short positions in firms with tangible collateral. I test whether the long-short portfolio has positive and significant abnormal return after controlling for conventional four factors (equation 2.2). If the abnormal return is positive and significant, the equity return difference between intangible collateral firms and tangible collateral firms (1) are not explained by conventional risk explanations, and (2) could imply collateralizability premium is not applied to this case.

Similar to the Fama-Macbeth regression, if collateralizability premium exists then the portfolio of firms with intangible collateral is expected to have lower stock returns due to the insurance effect of their intangible capital used as collateral. Because sample firms are mostly pledging intangible capital in addition to tangible capital as briefly mentioned in Introduction (Section 2.1), it is more reasonable to deem that the intangible collateral is related to ease of financial constraint at least in this sample.

$$r_{i,t} = \alpha_i + \beta_{1,i} MKTRF_t + \beta_{2,i} SMB_t + \beta_{3,i} HML_t + \beta_{4,i} UMD_t + \epsilon_{i,t} \tag{2.2}$$

I also add another evidence that shows that intangible collateral can relax financial constraint, I run a fixed effect panel regression of next year's firm-level leverage on intangible collateral usage. If intangible capital is related to ease of financial constraint, then future leverage should be higher when intangible capital is pledged. This regression setup rules out unobservable firm-level, industry-level, and year fixed effect.

$$Ln(Leverage_{i,t+1}) = const + \beta_{IP} IP_{i,t} + \Sigma \beta_{cont} Controls_{i,t} + FirmFE + YearFE + IndustryFE + \epsilon_{i,t} \tag{2.3}$$

For the next step, I perform stricter matching analysis that can provide a more accurate comparison to test the collateralizability premium hypothesis. I compare Cell(1) with Cell(4) in terms of stock returns after controlling for relevant covariates by nearest neighbor matching method.

The most basic assumption behind this setting is that intangible capital and tangible capital do not have any different nature when it comes to being pledged as collateral. In addition, if a matched pair of firms have almost same characteristics including intangible capital related variables but one pledges intangible capital while the other does not, then the stock return difference between the two is mostly due to the collateralizability of their intangible capital. In this case, the former one can benefit from the insurance effect as discussed in [Ai et al., 2018]. Therefore, the former should have lower stock returns than the latter. If it turns out to be opposite, then the collateralizability premium is not able to explain the stock return of firms pledging intangible capital as collateral.

I match treated firms that have syndicated loans backed by intangible collateral with other similar firms that do not have syndicated loans based on total asset size, intangible capital (standardized by assets), financial constraint (HP_index, net debt over sales, existence of credit ratings), intangible investment (R&D intensity measured by R&D over total assets), and access to public debt market.

I use nearest neighbor matching method that requires exact matches for indicator variables such as access-to-public-debt-market but closest matches for continuous variables. All matching dimensions have the same weights except the exact matching variables. I allow only one match for each treated observation (firm-year). Also, I exclude poorly matched pairs before doing analysis based on the condition that the calculated matching distance[1] between two firms should be lower than 1. This drops around 1.8 percent of the observations.

Once a matched sample is created, I form a long-short portfolio by taking long positions in treated firms and short positions in control firms. With the stock return series of the long-short portfolio, I run time-series regression with four factors to investigate the sign of the abnormal return of the portfolio.

In order to check if the full sample result is driven by a specific time period, I run sub-sample regressions. I first exclude the Dot-Com Crash period (the fiscal year 2000-2002), then early years that intangible collaterals are barely observed (before 2000), and finally financial crisis period (2007-2008).

Lastly, I calculate the intangible collateral portfolio's exposure to intangible related factors in order to briefly get a sense if there are any possible other factors that can explain the return difference between the intangible collateral portfolio and the tangible collateral portfolio/no syndicated loan portfolio. I estimate simple beta to knowledge capital productivity shock suggested in Chapter 1 and technology adoption shock suggested by [Lin et al., 2018]. The technology adoption shock carries a positive price of risk according to [Lin et al., 2018] and estimated from changes in technology standards. If the intangible collateral portfolio has nothing to do with this type of intangible related

[1] Euclidean norm is used.

factors, then no beta difference is expected.

2.4 Data and Variable Construction

I obtain loan-level collateral type data from *dealscan* database which covers 70 to 80 percent of US firms that files various filings to SEC. It has various information about each loan, such as borrower, lender, secured or not, collateral type, loan type, origination country, and so many other kinds of information. The variable *security* tells us the type of collateral a loan is backed by. Also, one can read comments associated with each loan and extract additional information about collateral types if the *security* variable is missing.

The sample period is from 1994 to 2015 because there is no record of intangibles used as collateral in Dealscan database before 1994. Firmyears with intangible collaterals are identified based on loan security information. If a firm uses intangible assets as collateral for a syndicated loan, all firmyears within the loan's maturity are considered as firmyears that are affected by intangible collaterals. To be more specific, Dealscan provides facility (tranche) level information. So I first identify loan facilities that use intangibles as collateral. I aggregate facilities over loan package level first and then borrower firmyear level, and give a value of 1 to firmyear level intangible collateral indicator variable. I only include corporate loans that have primary purpose[2] closely related to investment following [AlMEIDA and CAMPELLO, 2007]. Intangible assets include trademarks, brand names, customer lists, publishing titles, licensing agreements, franchise agreement, use rights, patented technology, software, trade secrets, and goodwill. I exclude facilities that pledge all assets of the borrower to reduce confounding effects from for example, M&A. My sample and sample in [Lumioti, 2012] are different because of different filtering. If I remove loan purpose filter and long term maturity filter, I could get a similar sample in [Lumioti, 2012]. My sample is more restricted to have proper comparability between intangibles-pledged loans and non-intangibles-pledged loans. Out of all facilities with non-missing security information, 12.1 percent of the facilities pledge intangibles as collateral.

From the sample of public firms, I exclude financial firms (sic code 6000 - 6999) and utility firms (sic code 4900-4999) from the sample following the convention. There are 1,341 publics firms in

[2]Acquisition line, Capital expenditure, Corporate purposes, and Takeover

the sample who use corporate loans to finance their investment-related projects between 1994 and 2015. By confining sample as above, I can control for borrower's credit-related variables as well as major loan characteristics that affect usage of corporate loan financing.

Among the sample firms, there are 116 firms (8.6 percent) that pledge intangibles to secure loans, which is small but not negligible. Most of the firms have loan packages that have both intangible secured facilities and other secured facilities. Therefore the intangible collateral effect found in this study is the effect of pledging intangible collateral in addition to conventional collateral.

All other financial data comes from COMPUSTAT and CRSP. Firm-level intangible capital is from [Peters and Taylor, 2016]. All the book values are based on fiscal year-end(December) values. Size is measured at the end of June, while Book-to-Market(BE/ME) is measured at the end of December, following conventions. The book value of equity is defined as shareholder equity(SEQ), plus deferred taxes($TXDITC$, if available), minus preferred stock($PSTK$). If SEQ is not available, I use common equity(CEQ) plus preferred equity($PSTK$) or Asset(AT) minus Liabilities(LT). If $PSTK$ is not available, I use redemption value($PSTKRV$) or liquidation value($PSTKL$). The market value of equity is measured at the end of the fiscal year(PRCC_F x CSHO). Net Debt is defined as total debt (DLTT + LCT) minus Cash and cash equivalents(CHE). Gross profitability is defined as gross profits(REVT minus COGS) divided by total assets. R&D/AT is R&D expenses(XRD) standardized by total assets. If R&D expenses are missing for firms that have any nonzero R&D expense in the past, I assign zero. Otherwise, I exclude the firm from the sample. Momentum is measured by multiplication of excess returns(raw) during the previous 12 months. Stock returns are matched with previous year's book values following Fama-Macbeth.

Net Secured Loan represents the total balance of principals of loans after amortization each firmyear. Average Maturity of Loan is a loan-size weighted average of remaining maturities of loans each firmyear. Following [AlMEIDA and CAMPELLO, 2007], I classify firms with credit ratings from AAA+ to BBB- as investment grade, from BB+ to S.D. as speculation grade based on COMPUSTAT S&P issuer long term credit ratings(splticrm). *Unrated* has a value of 1 if a firm doesn't have a long term credit rating.

The treated observations in the sample are restricted to firms with corporate loan sample, while control observations can be drawn from firms without corporate loans but with public debt. In order to avoid this to confound the results, I control for whether a firm has access to the public debt market. This information is first extracted from COMPUSTAT credit ratings. Also, I identify firm-years with notes and bonds from debt capital structure data of Capital IQ.[3] There is a variable telling type of debt instrument (CAPITALSTRUCTURESUBTYPEID), and it is coded as 4 for notes and bonds. If a firm has either credit rating or active notes and bonds, I give a value of 1 to the access-to-public-debt-market variable. Otherwise, I give a value of 0.

Lastly, to estimate technology adoption shock following [Lin et al., 2018], I obtain technology standards data from the Searle Center at Northwestern School of Law. The dataset contains technology standards that are active in each year from 1993. Each technology standards have a unique id so that one can count number of newly adopted standards every year. The technology adoption shock is logged growth rate of the number of unique standards in general technology fields including IT.

2.5 Findings and Discussions

Figure 2.1 shows fraction of intangible collateral usage from 1996 to 2014. While intangible capital was barely pledged in the '90s, it started being used as collateral from the beginning of the 2000s. The usage of intangible collateral peaked around the mid-2000s when the economy was in great condition, then started decreasing after the financial crisis and around the adoption of the Dodd-Frank Act in 2010. And it is still a relatively rare phenomenon. However, it is worthwhile to investigate its effect on firm variables and risks because this is a relatively new pattern that is not even assumed in many economic theories.

As shown in the Table 2.2, loan facilities that are secured by intangible collateral compose about 2 percent of the final sample. Intangibles-pledged facilities have higher LIBOR spread than non-intangibles-pledged facilities. The pooled mean difference is 73.52 basis points, which is statistically significantly different from zero. This could imply lenders may deem intangible collateral is not as safe as tangible collateral and require higher loan price. [Lumioti, 2012] also find intangible pledged loans have a higher spread. But there is no difference between the two groups in terms of the average maturity of loans or the average loan facility amount. Note that some of the intangible

[3]The data is provided via WRDS.

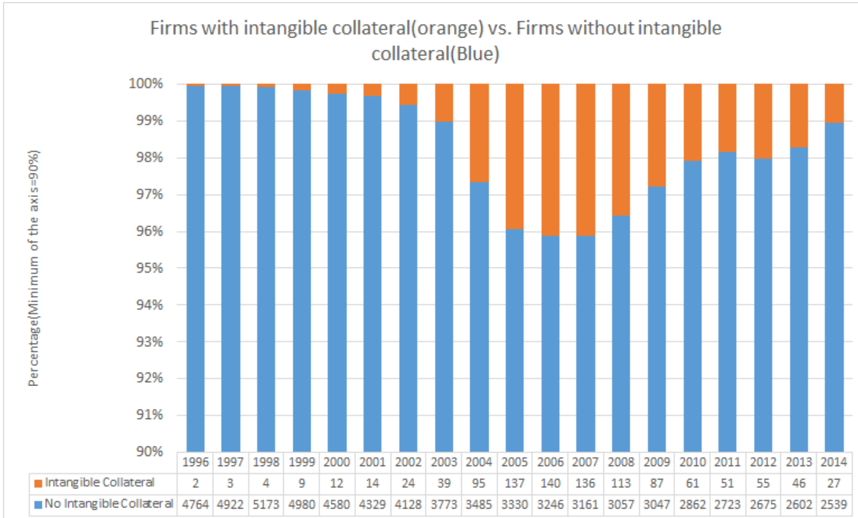

Figure 2.1: **Ratio of firms using intangible collateral by year**

pledged facilities in the sample could use other collaterals together. Therefore there is no significant difference in maturity or facility amount at the mean level. In terms of loan types, long term revolver/line is mostly found and then term loans. There is no noticeable difference in the types of loans between the two groups. More than 90 percent of facilities are syndicated. In the sample, 72 percent of non-intangibles pledged facilities are for general corporate purposes, while 60 percent of intangibles pledged facilities are for general purposes. 25 percent of intangibles pledged facilities are for takeover, while only 14 percent of non-intangibles pledged facilities are for takeover.

Table 2.3 Among all firms that are public, covered by Dealscan, and with secured long term corporate loans for investment purposes, 8.4 percent of firms have records of intangible collateral, and about 5 percent of firm-months are identified as intangible collateral months. Firms pledging intangibles as collateral are on average younger, smaller, but levered more, having bigger loans compared to their total debt size, getting one-month-longer term loans, doing 1.5 times more R&D, and slightly less profitable. The characteristics of intangible-collateral firms remind us of those of financially constrained firms discussed in the literature.

Table 2.4 displays credit ratings and financial constraint status of non-intangible-pledged firm-months and intangible-pledged firm-months. Following [Hadlock and Pierce, 2010], top tercile firms with high HP_index are classified as financially constrained, and bottom tercile firms with low HP_index are classified as financially unconstrained. Out of all unconstrained firm-months,

Table 2.2: **Loan Facility-level Summary Statistics**

This table contains summary statistics of loan facilities included in the sample. All loan facility data is obtained from LPC Dealscan. Sample period is from 1994 to 2014. Loans with primary purpose related to corporate investments are included. Only long term loans (maturity $>$ 12 months) are included. MinBps is the minimum number of basis points added to the current interest rate level to form the current facility interest rate or facility pricing. MaxBps is the maximum number of basis points added to the current interest rate level. Facilities using LIBOR as current interest rate level is included for statistics related to loan pricing. Facility amount is the size of loan facility.

Variable	N	Mean	Std Dev	Minimum	Maximum
		Panel A: Non Intangibles Pledge			
MinBps	21,800	207.496	140.580	0	1,500
MaxBps	21,832	207.542	140.547	0	1,500
Maturity (Months)	40,543	54.342	25.227	13	288
Facility Amount($Mil)	40,505	243.898	744.965	0	32,224.147
		Panel B: Intangibles Pledged			
MinBps	639	281.017	112.010	80	950
MaxBps	639	281.056	112.039	80	950
Maturity (Months)	780	55.867	20.390	13	182
Facility Amount($Mil)	780	252.744	657.968	0.223	9,750.000

Table 2.3: **Firm-level Summary Statistics**

This table summarizes firm level variables of non-intangible-pledged firm-months versus intangible-pledged firm-months. The sample period is from **1994** to **2014** for firm fiscal years, and from 1995 to 2016 for CRSP monthly returns.*Intangible Pledged* is an indicator variable that has value of 1 if a firmyear has secured corporate loan pledging intangibles as collateral. Firm level intangible capital is from [Peters and Taylor, 2016]. All the book values are based on fiscal year-end(December) values. Size is measured at the end of June, while Book-to-Market(BE/ME) is measured at the end of December, following conventions. The book value of equity is defined as shareholder equity(SEQ), plus deferred taxes($TXDITC$, if available), minus preferred stock($PSTK$). If SEQ is not available, I use common equity(CEQ) plus preferred equity($PSTK$) or Asset(AT) minus Liabilities(LT). If $PSTK$ is not available, I use redemption value($PSTKRV$) or liquidation value($PSTKL$). The market value of equity is measured at the end of fiscal year(PRCC_F x CSHO). Net Secured Loan represents total balance of principals of loans after amortization each firmyear. Average Maturity of Loan is loan-size weighted average of remaining maturities of loans each firmyear. Following [AlMEIDA and CAMPELLO, 2007], I classify firms with credit ratings from AAA+ to BBB- as investment grade, from BB+ to S.D. as speculation grade based on COMPUSTAT S&P issuer long term credit ratings(splticrm). *Unrated* has value of 1 if a firm doesn't have long term credit rating. Net Debt is defined as total debt (DLTT + LCT) minus Cash and cash equivalents(CHE). Gross profitability is defined as gross profits(REVT minus COGS) divided by total assets. R&D/AT is R&D expenses(XRD) standardized by total assets. If R&D expenses is missing for firms have nonzero R&D expense in the past, I assign zero. Momentum is measured by multiplication of excess returns(raw) during previous 12 months.

	N	Mean	S.D.	MIN	MAX
Panel A: Non-intangible collateral months					
Age	85,617	26.190	18.650	1.000	89
Assets	85,617	6,139.440	17,930.340	4.131	355,935
Size	85,617	6,361.050	14,364.600	5.069	76,688.570
BE/ME	85,617	0.567	0.448	0.050	3.303
Momentum	85,573	-0.987	0.067	-1.000	-0.349
Net Debt / Sale	85,509	0.381	0.469	-1.652	5.732
Debt / Assets	85,617	0.453	0.173	0.006	1.947
Net Secured Loan / Debt	85,617	0.707	0.794	0.004	5.518
Average Maturity	85,617	37.670	18.600	0.000	152
R&D / Assets	85,617	0.039	0.048	0.000	0.240
Gross Profitability	85,617	0.348	0.170	-0.073	1.084
Number of Firms	1,263				
Panel B: Intangible collateral months					
Age	4,486	19.609	16.757	1.000	65
Assets	4,486	2,094.630	3,678.357	7.932	25,458
Size	4,486	1,828.780	2,980.720	5.069	33,356.85
BE/ME	4,486	0.566	0.482	0.051	3.304
Momentum	4,474	-0.983	0.079	-1.000	-0.349
Net Debt / Sale	4,462	0.470	0.712	-1.652	5.732
Debt / Assets	4,486	0.497	0.171	0.082	0.946
Net Secured Loan / Debt	4,486	0.856	0.842	0.007	5.518
Average Maturity	4,486	38.210	17.608	0.000	84
R&D / Assets	4,486	0.061	0.067	0.000	0.240
Gross Profitability	4,486	0.320	0.183	-0.073	1.080
Number of Firms	116				

Table 2.4: **Summary Statistics: Financial Constraints**

This table summarizes credit status of non-intangible-pledged firm-months versus intangible-pledged firm-months. Following [AlMEIDA and CAMPELLO, 2007], I classify firms with credit ratings from AAA+ to BBB- as Investment grade, from BB+ to S.D. as Speculation grade based on COMPUSTAT S&P issuer long term credit ratings(splticrm). *Unrated* has value of 1 if a firm doesn't have long term credit rating. Constrained and Unconstrained firms are identified by Hadlock and Pierce's financial constraint index(below HP_index). HP_index is calculated by $-0.737Size + 0.043Size^2 - 0.040Age$. Size is adjusted by gdp deflator in 2009 dollars. Age is current year minus the very first year appearing in COMPUSTAT with non-missing stock price. Top tercile with high HP_index is considered constrained, while bottom tercile is considered unconstrained. All variables are indicator variables, thus mean value represents percentage of firm-months having value of 1 for each variable.

	Non-intangible firm-months		Intangible firm-months	
	count	mean	count	mean
Panel A: Credit Ratings				
Unrated	85,617	0.4914	4,486	0.4561
Speculation grade	85,617	0.3674	4,486	0.5439
Investment grade	85,617	0.1411	4,486	0.0000
Panel B: Financial Constraint				
Constrained	85,617	0.0998	4,486	0.2345
Unconstrained	85,617	0.5932	4,486	0.3629

only 3.2 percent is identified as intangible pledged firm-months, while 12.3 percent is identified as intangible pledged firm-months among the constrained sample. However, among intangible-pledged firm-months, 36.3 percent is unconstrained while 23.5 percent is constrained. Also, intangible firm-months are composed of more portion of speculation grade borrowers than non-intangible firm-months. 7.76 percent of speculation grade group is identified as intangible-collateral pledged. There is no intangibles-pledged firm-month in investment grade group and 4.86 percent of firm-months in the unrated group is identified as intangible-pledged firm-months. In sum, overall firms with intangible collateral are more likely to be facing tighter financial constraints and lower credit rating.

Table 2.5 contains Fama-Macbeth cross-sectional regression on $IP_{i,t}$ variable with other control variables. The results show that pledging intangibles could make cross-sectional differences in stock returns. Pledging intangibles as collateral increases stock returns by 0.9 percentage points per month, annualized 10.8 percentage points. According to [Ai et al., 2018], firms with a higher ratio of intangible capital should have higher stock returns than the other firms with a lower ratio of it, because intangible capital cannot provide insurance for economic downturns and tighter financial constraint. Af first glance, this result seems to support the collateralizability premium. However, the sample used in this study violates their assumption and mechanism because intangible capital is used as collateral and it might help relax financial constraints. If it turns out that intangible collateral relaxes financial constraints, then we would not expect to see higher stock returns for the firms pledging intangible collateral for corporate loans. Therefore, this evidence is interesting enough to motivate us to explore whether we can find empirical evidence consistent with/against the collateralizability premium with this specific sample.

In addition to the cross-sectional regression in Table 2.5, the results of time-series regression in Table 2.6 show that the firms with intangible collateral have not only higher average returns but also abnormal returns. The dependent variable is a long-short portfolio's return that takes long positions in intangible collateral stocks and short positions in the tangible collateral stocks. Even after controlling for the conventional four factors, the abnormal return of the intangible collateral portfolio is positive and significant. It is also an economically significant level of 9.6 percent to 12 percent annualized.

Table 2.7 shows that pledging intangible is positively associated with higher book leverage after controlling for conventional leverage determinants. If a firm has intangible collateral, the next year's leverage of the firm increases by about 10 percentage points evaluated at the mean logged leverage level (-2.3, 9.8% in raw value). One possible scenario for the firms with intangible capital is that, together with the result in Table 2.4, they currently face tighter financial constraint and use intangible capital as collateral to relax financial constraints so that their next year leverage increases.

Different from what we have typically assumed in the literature that only tangible capital is pledgeable and safe, the intangibles do play a role as collateral to secure corporate loans. Therefore this contradicts the first assumption of the collateralizability premium that intangible capital is not pledgeable. In addition, this provides support for performing another matching exercise to look at whether the higher returns of the firms with intangible collateral is against the collateralizability premium channel, because stock return should be lower for the firms who can benefit from the insurance effect of intangible collateral that can relax financial constraint and increase borrowing.

Table 2.5: **Fama-Macbeth Regression: Pledging intangibles and Cross-section of Stock Returns**

This table contains Fama-Macbeth regression of excess returns(percent) on intangibles used as collateral and various controls. The sample period is from **1994** to **2014** for firm fiscal years, and from 1995 to 2015 for CRSP monthly returns.*Intangible Pledged* is an indicator variable that has value of 1 if a firmyear has secured corporate loan pledging intangibles as collateral. Firm level intangible capital is from [Peters and Taylor, 2016]. Net secured loan represents total balance of principals of loans after amortization each firmyear. Average Maturity of Loan is loan-size weighted average of remaining maturities of loans each firmyear. Following [AlMEIDA and CAMPELLO, 2007], I classify firms with credit ratings from AAA+ to BBB- as investment grade, from BB+ to S.D. as speculation grade based on COMPUSTAT S&P issuer long term credit ratings(splticrm). *Unrated* has value of 1 if a firm doesn't have long term credit rating. Net Debt is defined as total debt (DLTT + LCT) minus Cash and cash equivalents(CHE). Knowledge Capital is total amount of intangible capital related to R&D and patents([Peters and Taylor, 2016]). Momentum is measured by multiplication of excess returns(raw) during previous 12 months (July of year t-1 to June of year t). 1 digit SIC code is used to control for industry fixed effect. All independent variables except indicator variables are winsorized at 1%.

VARIABLES	(1) Ex.Ret	(2) Ex.Ret	(3) Ex.Ret
Intangibles Pledged(IP)	0.823*	0.905**	0.935**
	(0.422)	(0.427)	(0.440)
Net Secured Loan/Debt			-0.00643
			(0.0717)
Average Maturity of Loan			0.223**
			(0.105)
Speculation Grade			-0.0317
			(0.194)
Unrated			0.0208
			(0.245)
Net Debt/Sale	-0.259	-0.196	-0.297
	(0.219)	(0.235)	(0.299)
Knowledge Capital/Assets		0.0866	0.0844
		(0.0890)	(0.0796)
R&D/Assets		0.0435	0.0553
		(0.0450)	(0.0401)
Gross Profitability	0.294*	0.173	0.238
	(0.169)	(0.188)	(0.167)
BE/ME	0.413**	0.501**	0.302**
	(0.168)	(0.212)	(0.120)
Size	0.000208	0.0188	-0.0369
	(0.0623)	(0.0656)	(0.0720)
Momentum	0.00175	0.00220	0.00905
	(0.0179)	(0.0176)	(0.0192)
Constant	2.138**	0.940	0.135
	(1.049)	(0.801)	(0.869)
Observations	91,126	88,281	85,400
R-squared	0.091	0.104	0.123
Number of months	264	264	264
Industry FE	YES	YES	YES

Standard errors in parentheses

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 2.6: **Time Series Regression: Intangible-pledging Portfolio minus Non-intangible-pledging Portfolio**

This table contains time series regression of excess return differences between intangible collateral firms and non-intangible collateral firms on conventional asset pricing factors. The sample period is from **1994** to **2014** for firm fiscal years, and from 1995 to 2016 for CRSP monthly returns. Portfolio returns are equally weighted.

VARIABLES	(1) Ret. Diff.	(2) Ret. Diff.	(3) Ret. Diff.
Alpha	1.011***	0.962**	0.808*
	(0.415)	(0.416)	(0.413)
MKT	-0.02498	-0.014	0.087
	(-0.090)	(-0.093)	(0.097)
SMB		0.044	-0.005
		(0.127)	(-0.126)
HML		0.199	0.290**
		(0.135)	(0.137)
UMD			0.239***
			(0.081)
R-Square	0.0003	0.0099	0.0472

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

If collateralizability premium works for the sample in this study, one can expect to see lower stock returns for the intangible collateral group than the tangible collateral group, taking into account the fact that intangible capital is pledged together with tangible capital most of the cases in the sample.

Accordingly, the next analysis tries to refine the sample by matching on several dimensions that are relevant to intangible collateral usage and a firm's financing ability and constraint. All the main results until here are done with the less sophisticated sample that compares a fairly small number of treated firms and a drastically larger number of control firms, even though several control variables are thrown in the Fama-Macbeth regression and Leverage regression. Therefore, as described in Section 2.3, I show the matched sample results in Table 2.8 and 2.9 that compare Cell (1) and Cell (4) of Table 2.1.

Table 2.8 shows averages of variables used in the matching analysis to compare Cell(1) and Cell(4) of Table 2.1. The treated firms are matched on asset size, financial constraint (HP Index), intangible capital standardized by assets, R&D intensity, and access to the public debt market. The

Table 2.7: **Panel Regression: Leverage**

This table contains fixed effect panel regression of book leverage on Intangible collateral indicator variable and other controls. Liquidity is defined as cash flow(EBIT) over assets(AT). Tangibility is estimated following [AlMEIDA and CAMPELLO, 2007], CHE + 0.715RECT +0.547INVT + 0.535PPENT. For the rest of variables please see: Table 2.5. The leverage is in log value and so other control variables are except indicator variables (Unrated, Speculation grade).

VARIABLES	(1) Leverage
Intangibles Pledged (IP)	0.602**
	(0.277)
Liquidity	-0.194
	(0.206)
Tangibility	-2.324***
	(0.447)
ROA	-0.453
	(0.316)
Size	-0.354***
	(0.123)
BE/ME	-0.443***
	(0.111)
Unrated	-1.252***
	(0.321)
Speculation Grade	0.102
	(0.176)
Constant	-2.541**
	(1.121)
Observations	84,157
R-squared	0.635
Firm FE, Year FE, Industry FE	YES
Clustered	Firm

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 2.8: **Matched Sample Summary Stats**

	Control	Treat
Assets	2378.274	2563.971
HP Index	-3.517	-3.525
Intangible Capital / AT	0.596	0.607
Net Debt / SALE	0.319	0.480
R&D / AT	0.030	0.031
Avg. No. Firms (month)	42.330	

nearest neighbor matching results return average distance of 0.126, and all the dimensions except for net debt over sales are very closely matched. Due to the small sample size of treated firms and one-to-one matching method, the average number of firms per month is around 42. With this matched sample, I form treated and control portfolios. Then I form a long-short portfolio by long the treated portfolio and short the control portfolio.

Table 2.9 shows if this long-short portfolio shows any significant portion of the returns that are not explained by conventional asset pricing factors. Column (1) of Table 2.9 is a full-sample time series regression result, and it is clearly shown that α is positive and significant, with an annualized return of 6 percent. Therefore, it is plausible that the treated firms, which have syndicated loans backed by intangible collateral, have higher stock returns than other similar firms in terms of intangible capital, intangible investment, size, and financial constraint. This could be evidence against the collateralizability premium. The matched firms have similar intangible capital and investment and financial constraint, but they can not pledge it to borrow using syndicated loans. The collateralizability premium story should predict that these matched firms should have higher stock returns because they lack collateralizable intangible capital even if they have similar amount of intangible capital.

In order to investigate if the result is driven by some specific time period, I plot time-series returns of the long-short portfolio as in Figure 2.2. Because of the small number of observations before the year 2000, the graph is very volatile. Other than this, the Dot-com Crash period and Financial Crisis period also show relatively more volatile patterns. Therefore in Column(2), (3), and (4), I run the same time series regression excluding the specific periods that might solely drive the

Table 2.9: **Time Series Regression: Firms with Syndicated loans backed by Intangible Collateral minus Matched Firms without Syndicated Loans**

This table contains time series regression of excess return differences between intangible collateral firms and non-intangible collateral firms on conventional asset pricing factors. The sample period is from 1994 to 2014 for firm fiscal years, and from 1995 to 2016 for CRSP monthly returns. Portfolio returns are equally weighted. Column (1) contains full sample results. Column (2) excludes Dot-com crash period in early 2000s from 2000 to 2002. Column (3) excludes years before 2000 in which data is very noisy due to low number of observations. Column (4) excludes financial crisis period from 2008 to 2009.

VARIABLES	(1) Ret. Diff.	(2) Ret Diff	(3) Ret Diff	(4) Ret Diff
α	0.501***	0.551***	0.441***	0.474***
	(0.0938)	(0.0986)	(0.105)	(0.102)
MktRF	0.0127	0.00661	-0.00327	0.0278
	(0.0238)	(0.0256)	(0.0299)	(0.0273)
SMB	0.0266	0.0363	0.118**	0.00536
	(0.0425)	(0.0479)	(0.0474)	(0.0412)
HML	0.113***	0.115***	0.0509	0.121***
	(0.0341)	(0.0381)	(0.0411)	(0.0417)
UMD	0.0810**	0.0627	0.0995**	0.103***
	(0.0321)	(0.0383)	(0.0454)	(0.0295)
Observations	252	216	180	228
R-squared	0.107	0.057	0.157	0.126

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

result. Column (2) excludes Dot-com crash period in the early 2000s from 2000 to 2002. Dot-com crash is somewhat important for the result because the r-squared is slightly higher than half of the r-squared in full sample analysis (column(1)). Although r-squared decreases by half compared to the r-squared of column(1), the abnormal return of column(2) is greater than that of column(1) and still statistically significant. Column (3) excludes years before 2000 in which data is very noisy due to the low number of observations. Column (4) excludes financial crisis period from 2008 to 2009. Similar to the result of Column (2), the abnormal returns are positive and significant for the sub-periods in Column (3) and (4). Of course the magnitude of α decreases by 0.026 percentage points per month (Column (1) vs. Column (4)), this is only around 0.36 percentage point of yearly return. Therefore, it is shown that the abnormal return of the long-short portfolio is not just driven by several specific periods, and one can conclude that this could be empirical evidence against the collateralizability premium story.

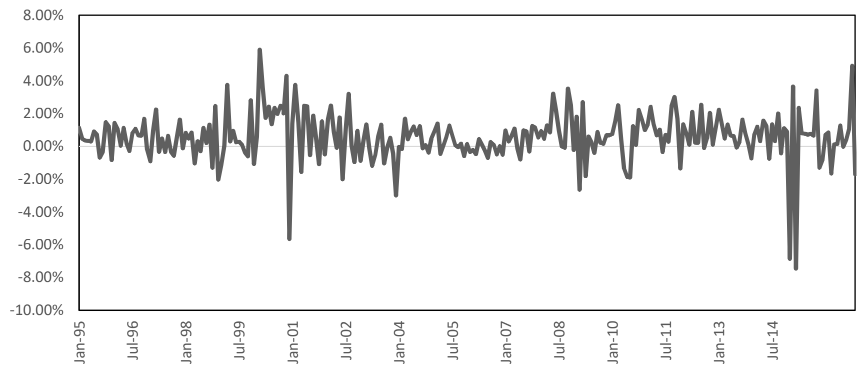

Figure 2.2: **Long-Short Portfolio's Monthly Returns over Time**
This figure is a time series plot of stock returns of firms having syndicated loan back by intangible collateral minus stock returns of matched firms not having syndicated loan. Firms are matched by asset size, R&D investments, financial constraint measured by HP index, public debt market access, and intangible capital standardized by total assets.

Then why do the firms with intangible collateral have higher stock return even with the ability to relax financial constraint? Although this is beyond the scope of this paper and should be studied in a separate paper, I provide a suggestive hypothesis briefly before I wrap up the discussion of this chapter. In connection with Chapter 1, the firms with intangible capital could possibly be more exposed to any risk sources that are closely related to intangible investment, productivity, or aggregate technology adoption shock. Hence I calculate sensitivities of the treated and control portfolios to knowledge capital productivity shock suggested in Chapter 1 and technology adoption shock following [Lin et al., 2018]. Note that technology adoption shock beta is estimated from 20-year time series data at an annual frequency, so the beta might not be very stable.

Table 2.10 shows that the treated firms have higher betas than the control firms. One can carefully conjecture that the firms with intangible collateral could be more exposed to intangible or technology related risks by looking at the difference in betas. It would be interesting to study why intangible collaterals should be related to technology risk or intangible productivity risk. Of course, there is still a possibility that collateralizability premium can explain the findings in this paper. For example, if intangible collateral relaxes financial constraint less than tangible collateral does, then the benefit from the insurance effect of the collateral is smaller for firms with intangible collateral than firms with tangible collateral. In this case, firms with intangible collateral could have

Table 2.10: **Exposure to Technology Adoption S hock**

This table contains betas of treated portfolio and matched portfolio to aggregate technology adoption shock and knowledge capital productivity shock. The technology adoption shock is estimated following [Lin et al., 2018]. Knowledge capital productivity shock is estimated following Section 1.3. The treated portfolio contains firms with syndicated loans backed by intangible collateral. The control portfolio in Panel A contains firms without syndicated loans. The control portfolio in Panel B contains firms with only tangible collateral.

	β_{know}	β_{tech}
	Panel A: Cell(1) vs. Cell(4) in Table 2.1	
Treat	0.93	0.08
Control	0.52	0.04
	Panel A: Cell(1) vs. Cell(2) in Table 2.1	
Treat	0.94	0.08
Control	0.31	-0.05

higher stock returns than firms with tangible collateral by the mechanism of the collateralizability premium model. However, the deeper discussion on why the firms with intangible collateral have higher stock returns than firms with tangible collateral only is off from the central purpose of this essay, and I leave it for future study.

2.6 Conclusion

By using LPC Dealscan data, I test the collateralizability premium hypothesis to explain higher stock returns of firms with intangible collateral. In the motivating analysis, I find statistically and economically significant stock returns for firms with intangible collateral than firms with only tangible collateral. The long-short portfolio has positive abnormal returns around 8 to 10 percent (annualized) after controlling for conventional factors.

Together with the results, I find evidence that intangible collateral can relax financial constraint by the panel regression of leverage on Intangible collateral indicator variable. The firms with intangible collateral achieve similar or slightly higher level of leverage compared to the other firms with only tangible collateral.

These results act as counter-evidence for the collateralizability premium hypothesis. What the evidence tells us is that the Sample firms can clearly relax financial constraint with intangible collateral, but they still earn higher abnormal returns. This is against the mechanism of the collateralizability premium that firms with higher collateralizability should earn lower stock returns.

In order to support my findings more, I perform matching exercises and compare the stock returns of firms with syndicated loans backed by intangible collateral and other firms without the syndicated loans even with the same level of intangible capital and R&D investments. The firms with syndicated loans backed by intangible collateral are expected to have lower stock returns than the matched firms without syndicated loans because they are more likely to have intangible capital with higher collateralizability.

Even with the nearest neighbor matching exercise, I find evidence against the collateralizability premium hypothesis that firms with intangible collateral still have significantly higher abnormal returns around 6 percent per year. The results are robust over several subsample periods.

In sum, even if firms have higher collateralizability, it does not have lower stock returns. Of course, the findings in this study do not fully rule out the existence of the collateralizability premium due to the small sample problem, one possible conclusion that can be drawn from the findings is that the collateralizability premium in [Ai et al., 2018] might be confounded by the effect of intangible capital which is independent with financial constraint.

Lastly, I show that intangible collateral portfolio has higher betas to technology-related shocks. Thus, it might be the case that firms with intangible collateral have higher stock returns due to its higher exposure to aggregate technology risk. This topic should be studied more deeply in the future.

References

[Ahuja and Lampert, 2001] Ahuja, G. and Lampert, C. M. (2001). Entrepreneurship in the large corporation: a longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal*, 22(6-7):521–543.

[Ai et al., 2012] Ai, H., Croce, M. M., and Li, K. (2012). Toward a Quantitative General Equilibrium Asset Pricing Model with Intangible Capital. *The Review of Financial Studies*, 26(2):491–530.

[Ai et al., 2018] Ai, H., Li, J., Li, K., and Schlag, C. (2018). The collateralizability permium. *Working Paper*.

[AlMEIDA and CAMPELLO, 2007] AlMEIDA, H. and CAMPELLO, M. (2007). Financial constraints, asset tangibility, and corporate investment. *Review of Financial Studies*, 20(5):1429–1460.

[Belo et al., 2014] Belo, F., Lin, X., and Vitorino, M. A. (2014). Brand capital and firm value. *Review of Economic Dynamics*, 17(1):150 – 169.

[Berk et al., 2004] Berk, J. B., Green, R. C., and Naik, V. (2004). Valuation and return dynamics of new ventures. *The Review of Financial Studies*, 17(1):1–35.

[Chan et al., 2002] Chan, L. K. C., Lakonishok, J., and Sougiannis, T. (2002). The stock market valuation of research and development expenditures. *The Journal of Finance*, 56(6):2431–2456.

[Corrado and Hulten, 2014] Corrado, C. A. and Hulten, C. R. (2014). *Innovation Accounting*, pages 595–628. University of Chicago Press.

[Dahlin and Behrens, 2005] Dahlin, K. B. and Behrens, D. M. (2005). When is an invention really radical?: Defining and measuring technological radicalness. *Research Policy*, 34(5):717–737.

[Dalle et al., 2017] Dalle, J.-M., den Besten, M., and Menon, C. (2017). Using crunchbase for economic and managerial research. *OECD Science, Technology and Industry Working Papers*.

[Fleming, 2001] Fleming, L. (2001). Recombinant uncertainty in technological search. *Management Science*, 47(1):117–132.

[Fontana et al., 2013] Fontana, R., Nuvolari, A., Shimizu, H., and Vezzulli, A. (2013). Reassessing patent propensity: Evidence from a dataset of r&d awards, 1977-2004. *Research Policy*, 42:1780–1792.

[Hadlock and Pierce, 2010] Hadlock, C. J. and Pierce, J. R. (2010). New Evidence on Measuring Financial Constraints: Moving Beyond the KZ Index. *The Review of Financial Studies*, 23(5):1909–1940.

[HAHN and LEE, 2009] HAHN, J. and LEE, H. (2009). Financial constraints, debt capacity, and the cross-section of stock returns. *The Journal of Finance*, 64(2):891–921.

[Hirshleifer et al., 2013] Hirshleifer, D., Hsu, P.-H., and Li, D. (2013). Innovative efficiency and stock returns. *Journal of Financial Economics*, 107(3):632 - 654.

[Hirshleifer et al., 2018] Hirshleifer, D., Hsu, P.-H., and Li, D. (2018). Innovative originality, profitability, and stock returns. *The Review of Financial Studies*, 31(7):2553–2605.

[Hochberg et al., 2018] Hochberg, Y. V., Serrano, C. J., and Ziedonis, R. H. (2018). Patent collateral, investor commitment, and the market for venture lending. *Journal of Financial Economics*, 130(1):74 – 94.

[Hopp et al., 2018] Hopp, C., Antons, D., Kaminski, J., and Salge, T. O. (2018). Disruptive and radical innovation.

[Hsu, 2009] Hsu, P.-H. (2009). Technological innovations and aggregate risk premiums. *Journal of Financial Economics*, 94(2):264 – 279.

[Jaffe, 1989] Jaffe, A. B. (1989). Characterizing the "technological position" of firms, with application to quantifying technological opportunity and research spillovers. *Research Policy*, 18(2):87 – 97.

[Jaffe and de Rassenfosse,] Jaffe, A. B. and de Rassenfosse, G. Patent citation data in social science research: Overview and best practices. *Journal of the Association for Information Science and Technology*, 68(6):1360–1374.

[Kaplan and Vakili, 2015] Kaplan, S. and Vakili, K. (2015). The double-edged sword of recombination in breakthrough innovation. *Strategic Management Journal*, 36(10):1435–1457.

[Kogan et al., 2017] Kogan, L., Papanikolaou, D., Seru, A., and Stoffman, N. (2017). Technological innovation, resource allocation, and growth*. *The Quarterly Journal of Economics*, 132(2):665–712.

[Lamont et al., 2015] Lamont, O., Polk, C., and Saaá-Requejo, J. (2015). Financial Constraints and Stock Returns. *The Review of Financial Studies*, 14(2):529–554.

[Lee et al., 2018] Lee, C. M., Sun, S. T., Wang, R., and Zhang, R. (2018). Technological links and predictable returns. *Journal of Financial Economics*, forthcoming.

[Li et al., 2014] Li, E. X. N., Liu, L. X., and Xue, C. (2014). Intangible assets and cross-sectional stock returns: Evidence from structural estimation. *Working Paper*.

[Lin et al., 2018] Lin, X., Palazzo, B., and Yang, F. (2018). The risks of old capital age: Assest pricing implications of technology adoption. *Working Paper*.

[LIVDAN et al., 2009] LIVDAN, D., SAPRIZA, H., and ZHANG, L. (2009). Financially constrained stock returns. *The Journal of Finance*, 64(4):1827–1862.

[Lumioti, 2012] Lumioti, M. (2012). The use of intangible assets as loan collateral. *SSRN Working Paper*.

[Mann, 2018] Mann, W. (2018). Creditor rights and innovation: Evidence from patent collateral. *Journal of Financial Economics*, 130(1):25 – 47.

[Novy-Marx, 2011] Novy-Marx, R. (2011). Operating leverage. *Review of Finance*, 15(1):103–134.

[NovyMarx, 2013] NovyMarx, R. (2013). The other side of value: The gross profitability premium. *Journal of Financial Economics*, 108(1):1–28.

[Peters and Taylor, 2016] Peters, R. H. and Taylor, L. A. (2016). Intangible capital and the investment-q relation. *Journal of Financial Economics*, Forthcoming.

[Trajtenberg et al., 1997] Trajtenberg, M., Henderson, R., and Jaffe, A. (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, 5(1):19–50.

[Turan G. Bali, 2016] Turan G. Bali, Robert F. Engle, S. M. (2016). *Empirical Asset Pricing: The Cross Section of Stock Returns*. John Wiley & Sons.

[U.S. Patent Office, 2012] U.S. Patent Office (2012). Overview of the u.s. patent classification system (uspc).

[Uzzi et al., 2013] Uzzi, B., Mukherjee, S., Stringer, M., and Jones, B. (2013). Atypical combinations and scientific impact. *Science*, 342(6157):468–472.

[Verhoeven et al., 2016] Verhoeven, D., Bakker, J., and Veugelers, R. (2016). Measuring technological novelty with patentbased indicators. *Research Policy*, 45(3):707–723.

[Wang, 2015] Wang, X. (2015). Catering innovation: Entrepreneurship and the acquisition market. *Working paper*.

[Whited and Wu, 2006] Whited, T. M. and Wu, G. (2006). Financial constraints risk. *The Review of Financial Studies*, 19(2):531–559.

[Wolf, 1996] Wolf, G. (1996). Steve jobs: The next insanely great thing. *Wired*.

[Yan and Luo, 2017] Yan, B. and Luo, J. (2017). Measuring technological distance for patent mapping. *Journal of the Association for Information Science and Technology*, 68(2):423–437.

[Younge and Kuhn, 2016] Younge, K. and Kuhn, J. (2016). Patenttopatent similarity: a vector space model. *SSRN*, Working Paper.